THE PROCEEDINGS OF THE PHYSICAL SOCIETY

Section B

Vol. 63, Part 9 1 September 1950 No. 369 B

Voltage Gradients in Long Gaseous Spark Channels

By J. B. HIGHAM and J. M. MEEK

Department of Electrical Engineering, University of Liverpool

MS. received 9th March 1950

ABSTRACT. Oscillographic measurements have been made of the variation with time of the voltage drops in long spark channels conducting aperiodic impulse currents rising to a peak in about 1/4 μsec. and decaying to half-value in from 10 to 85 μsec., with peak currents ranging from 60 to 700 amp. The lengths of the spark channels have been measured photographically, and the mean voltage gradients along the channels have been estimated. Results are given for sparks in air, nitrogen, oxygen and hydrogen at atmospheric and reduced pressures. Some preliminary measurements are described for sparks in water and oil, and for sparks in glass tubes. The voltage gradients in the various gases, with the exception of hydrogen, are independent of the peak current, but are influenced by the rate of current decay. The results are correlated with the current densities in spark channels, as recorded in a separate investigation.

§ 1. INTRODUCTION

The growth of a spark is initiated by a streamer process, sometimes termed a leader stroke, which develops across the gap to form a conducting channel bridging the two electrodes. The leader stroke is followed by a main stroke, sometimes termed a return stroke, which completes the breakdown process, and the external circuit then discharges through the pre-ionized channel (Allibone and Meek 1938).

Before the development of the leader stroke the voltage across the spark gap has a high value, equal to or greater than the minimum breakdown voltage for the gap. On the occurrence of the main stroke the voltage falls to a low value which corresponds to the voltage drop along the spark channel itself. The purpose of the present investigation has been to measure this voltage drop and to determine its variation with time for different gaps in several gases under various circuit conditions. Some preliminary results have been referred to in an earlier publication (Meek 1947).

Previous experiments for gaps up to 1 cm. in length have been made by Craggs and Meek (1946), who showed that the voltage drop along sparks carrying 120 amp. for 3·6 μsec. in argon and in hydrogen at atmospheric pressure averages roughly 90 volts per cm.

In two oscillograms published by Flowers (1943), the voltage gradient for spark channels in air across a gap of 8·8 cm. is seen to be about 500 volt/cm. at a time of 1 μsec. after breakdown, and before the currrent has reached its peak amplitude. Also, in measurements by McCann and Clark (1943) of the dielectric recovery characteristics of long air gaps, voltage gradients of 200 to 300 volts/cm. were recorded for sparks occurring along pre-ionized channels formed by earlier sparks. Norinder and Karsten (1949) record voltage gradients of the order of 100 volt/cm. in spark channels conducting oscillatory currents.

A complication in the earlier measurements with gaps of 1 cm. or less is that the voltage drops at the electrodes are comparable with the voltage drop in the discharge channel. A further complication is the diffusion of metal vapour into the gap, at a speed of about 10^5 cm/sec., so that the whole of the spark channel between electrodes spaced at 1 cm. contains both gas and metal vapour at times greater than about 5 μsec. from its initiation (Williams, Craggs and Hopwood 1949, Llewellyn-Jones 1945, 1946) and the voltage drop may therefore differ from that for a spark channel in the gas alone. Both these effects are of decreasing importance as the gap length is increased.

In the present investigation an attempt has been made to avoid the complications due to the electrodes by making measurements for longer gaps of length from 10 cm. up to 40 cm. Various gases have been studied, including air, nitrogen, oxygen and hydrogen at atmospheric and reduced pressures, and some preliminary measurements have been made for sparks in liquids with a shorter gap length.

§ 2. MEASUREMENT TECHNIQUE

Single impulse voltages were applied to the test spark gap from an 8-stage 1,000 kv. impulse generator of discharge capacitance 0·025 μF. The spark channel currents used had peak magnitudes between 60 and 700 amp., with various rates of decay determined by the series resistor R_s, as shown in Figure 1. The currents reached their peak values in approximately 0·25 μsec., and the values of R_s were chosen so that in general the currents decayed to half value in either 10 or 28 μsec. Voltages across the channels were measured at times ranging between 0·25 and 80 μsec. after breakdown. The times interval between successive sparks was so long that each spark was uninfluenced by the preceding one.

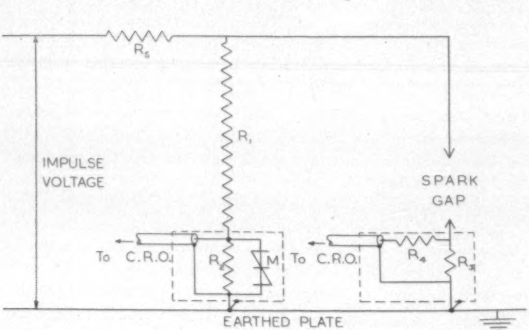


Figure 1. Circuit for voltage and current measurements.

Oscillographic measurements of voltage and current were made with a 10 kv. sealed-off cathode-ray oscillograph (Prime and Saxe 1949) the screen of which was photographed on 35 mm. film with a 2-inch f/1·0 bloomed lens. Time sweep durations down to 0·5 μsec. were available.

In order that the voltage across the spark gap could be recorded on the oscillograph, a potential divider had to be used, the dividing ratio of which was governed by the consideration that the voltage drop across the spark channel had to produce a measurable deflection on the oscillograph. If this is done with a normal resistance divider, the voltage across the gap before the gap breaks down is so high that it can cause damage to the oscillograph. For this reason a potential divider was constructed as shown schematically in Figure 1, with a non-linear resistor shunting the lower resistor R_2 of the divider.

Both high and low voltage resistance elements R_1 and R_2 were made from resistance ribbon (in which the warp is of silk and the weft of constantan wire) which has a reasonably low uniform inductance of about 1 μH. per inch. R_2 had a value of about 65 ohms, matching the surge impedance of the 12 ft. long polythene insulated concentric cable enclosed in conduit leading to the oscillograph. Shunting R_2 was a disc of Metrosil non-linear resistance material, the characteristic of which may be expressed in the form $V = KI^{\beta}$ where V = voltage, I = current, K = a constant, and the index β is about 0·25. The sample used was so chosen that its effective resistance is large compared with R_2 when the comparatively low post-breakdown voltage across a spark channel is being recorded; thus the divider is linear in its response to such voltages, but when the pre-breakdown voltage is applied, the effective resistance of the Metrosil disc is low enough to limit the divider output voltage to a safe value. Metrosil responds to transients without appreciable time lag, a consideration which precludes the use of a glow-tube or a similar safety device. Tests were made with and without the semiconductor, which showed that its presence did not effect the response of the potential divider to voltage transients similar to those being measured. A range of low value 'non-inductive' resistance units, R_3, were available for insertion in the circuit when measurements of current were required. A typical oscillogram of the residual voltage drop across a spark channel air is shown in Figure 2 (Plate I).

The usual precautions adopted in high-voltage measurements were adhered to throughout the work, and at least eight oscillograms were taken for any one condition being investigated. The oscillograms were then enlarged by about eight diameters, and tracings were made which were analysed with a specially photographed graticule. It was found that the axes of the cathode-ray oscillograph were 3° from normal, and the graticule was constructed accordingly. The zero for measurements of time from breakdown was taken where the rapidly decaying initial portion of the voltage oscillogram is asymptotic to the Y axis of the graticule, as illustrated in Figure 2. The possible error involved is not more than about 0·1 μsec.

§ 3. SPARK LENGTHS

The voltage drops measured across successive spark channels were found to have a 'scatter' not accountable for by measurement errors. It was realized early in the investigation that this was probably caused by variations in channel lengths, and consequently an attempt was made to record these lengths.

The path of any particular spark channel can be determined by taking photographs from two directions at right angles, but the subsequent analysis involved is prohibitive for a large number of records. A fairly close estimate of the total length may be obtained on a statistical basis from photographs taken with one camera, and this procedure has been adopted in the present investigation.

The estimated mean total spark length, L, may be expressed as

$$L = L_g + k(L_p - L_g),$$

where L_g is the gap length, L_p the mean photographed or projected length, and k a factor to be determined.

From a consideration of hypothetical spark formations, such as semicircular in a plane and helical in three dimensions, and allowing for small irregularities, the authors have assigned the value 2·0 to the factor k. However, the value of k is not critical; for example, if the mean photographed length is 10% greater than the gap length and the correct value of k is 2·0, the error involved by taking values of $k = 2{\cdot}3$ and 1·7 is only $\pm 2\frac{1}{2}\%$ of the correct length.

Photographs were taken of up to 50 sparks under each set of conditions of gap length, circuit constants, current, gas and gas pressure employed in the voltage measurements. Typical photographs are shown in Figure 3 (Plate II), where the changes in spark length for the different conditions are readily apparent. The spark channels in hydrogen are particularly irregular, and, for an applied voltage of approximately twice the minimum breakdown voltage, the mean total spark length is 68% greater than the gap length. The corresponding value measured for air is only 10%. For other conditions investigated, the lengths measured in hydrogen were more than in the other gases but not by so great an amount.

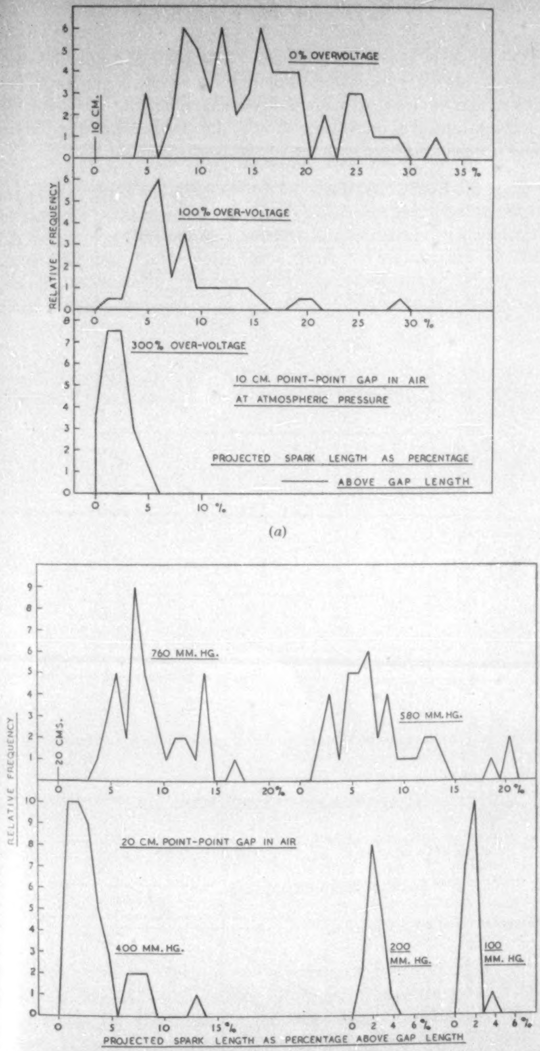
In Figures 4(*a*) and 4(*b*), the photographed spark lengths, expressed as a percentage above the gap lengths, are plotted against their frequency of occurrence for particular conditions. The resutant distribution curves illustrate the general tendency for the spark lengths to dminish and to become more uniform as the applied voltage is increased, or as the gas pressure is reduced with the applied voltage kept constant.

This is clearly shown in the lower graph of Figure. 5, in which the mean total spark length is plotted as a function of gap length for the case where the applied voltage was directly proportional to the gap length. The two curves correspond to the first two columns of Table 1 (see §4.1). In curve A, the voltage applied to the 10 cm. gap was little more than that required for breakdown, and the resulting mean spark length is high. As the minimum breakdown voltage for long point-point gaps does not increase linearly with gap length, the gaps longer than 10 cm. were subjected to voltages appreciably higher than the minimum breakdown voltages. Thus the curves bear out the general finding that the spark length diminishes with increasing over-voltage.

§4. VOLTAGE GRADIENTS

For any one set of conditions being investigated, several oscillograms have been obtained to show the total voltage drop across the spark gap as a function of time, and from the analysis of these oscillograms a mean curve has been drawn showing the variation with time of the ratio of the mean voltage drop to the mean measured length L of the spark. This curve may then be considered to show the mean voltage gradient along the spark channel as a function of time.

This method of determining the voltage gradient rests on the assumption that the gradient is uniform along the spark channel, neglecting electrode effects, and



(*a*)

(*b*)

Figure 4. Graphs showing the relative frequency of occurrence of spark lengths, expressed as a percentage above the gap length:

(*a*) For various applied voltages at constant gas pressure.

(*b*) For various gas pressures at constant applied voltage.

seems to be justified by the results obtained. The graph of Figure 6, in which the mean voltage across the spark channel is plotted as function of gap length at various times from breakdown for a particular current characteristic, illustrates this linear relation and indicates that the voltage drop at the electrodes and in the electrode connections is small compared with the voltage drop across the spark.

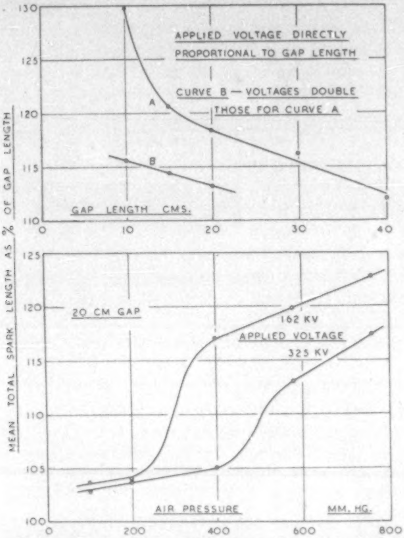


Figure 5. The dependence of spark length on applied voltage and gas pressure.

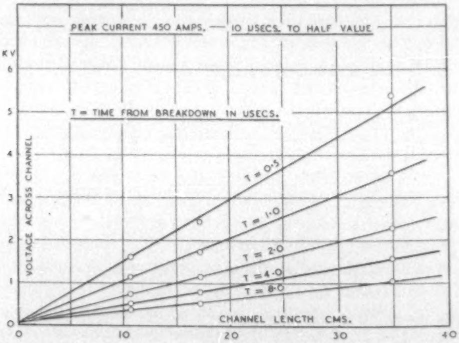


Figure 6. The linear relation between voltage drop and channel length at various times from breakdown.

As a further check on the effect of electrode connections, voltage oscillograms were recorded with the usual potential divider with the spark gap short circuited. Apart from small initial oscillations dying out in less than 0·5 μsec. the result was a straight trace coincident with the zero line.

4.1. *Air at Atmospheric Pressure*

Some of the results obtained are shown in Tables 1 and 2 and apply to spark channels between point electrodes, one of which is earthed, when the high voltage electrode is subjected to impulse voltages of positive polarity through series resistors of 750 and 2,000 ohms respectively. The currents are the values corresponding to the related sets of voltage measurements at various times from breakdown. These are taken from current oscillograms such as those reproduced in Figure 7 (Plate I). Columns (3), (6) and (9) give the mean total voltage measured across the sparks at various times from breakdown. The adjacent columns (4), (7) and (10) give the mean voltage gradient along the spark channel obtained by dividing the kilovolt values by the corresponding mean spark lengths.

Table 1

(1) Gap (cm.)	(2) μsec. *	(3) kv.	(4) v/cm.	(5) amp.	(6) kv.	(7) v/cm.	(8) amp.	(9) kv.	(10) v/cm.	(11) amp.
10	0·5	2·15	165	150	1·63	142	300	1·62	152	450
	1	1·41	108	145	1·15	100	290	1·15	107	436
	2	0·95	73	135	0·84	73	270	0·75	70	406
	4	0·60	46	118	0·60	52	236	0·53	49·5	350
	8	0·45	34·5	89	0·42	36·5	178	0·39	36·5	266
	16	0·32	24·5	50	0·28	24·5	100			
	24	0·24	18	29	0·20	17·5	58			
15	0·5	3·11	173	225	2·43	142	450	2·70	164	675
	1	2·01	112	218	1·72	100	436	1·80	109	654
	2	1·30	72	203	1·15	67	406	1·14	69	609
	4	0·85	47	175	0·79	46	350	0·82	50	525
	8	0·60	33	133	0·54	31·5	266			
	16	0·45	25	75	0·38	22	150			
	24	0·32	17·5	44	0·26	15	88			
20	0·5	3·54	150	300	3·14	139	600			
	1	2·38	101	290	2·18	96	580			
	2	1·60	68	270	1·50	66	540			
	4	1·14	48	236	1·06	47	472			
	8	0·78	33	178	0·76	33·5	356			
	16	0·56	24	100	0·53	23·5	200			
	24	0·38	16	58	0·34	15	116			
30	0·5	5·40	155	450						
	1	3·60	103	436						
	2	2·31	66	406						
	4	1·60	47	350						
	8	1·08	31	266						
	16	0·78	22·5	150						
	24	0·52	15	88						
40	0·5	6·48	145	600						
	1	4·64	104	580						
	2	3·00	67	540						
	4	2·04	45·5	473						
	8	1·32	29·5	356						

0·025 μF. impulse generator.
Series resistance 750 ohms.
Air at atmospheric pressure.
Current decaying to half value in 10 μsec.

* Column (2) gives time from breakdown.

For either current waveform there is close agreement in the magnitude of the voltage gradients for the various gap lengths and peak currents. For example, referring to Table 1, for gap lengths of 10 to 40 cm. and peak currents of 150 to

675 amp. the voltage gradients obtained 2 μsec. after breakdown lie in the range of 66 to 73 volt/cm. This is almost certainly within the limits of overall experimental accuracy.

A comparison of Tables 1 and 2 indicates the general effect of current duration. This effect is shown more clearly in Figure 8 in which voltage gradients for various current durations are plotted. The lower the rate of current decay, the higher the voltage maintained across the spark channel.

Table 2

(1) Gap (cm.)	(2) μsec.*	(3) kv.	(4) v/cm.	(5) amp.	(6) kv.	(7) v/cm.	(8) amp.	(9) kv.	(10) v/cm.	(11) amp.
10	0·5	1·80	145	60	1·45	130	125	1·62	151	190
	1	1·37	110	59	1·13	101	124	1·17	109	188
	2	1·04	84	58	0·87	78	121	0·83	78	183
	4	0·78	63	55	0·62	55·5	115	0·61	57	175
	8	0·54	43·5	50	0·48	43	105	0·49	46	160
	16				0·37	33	86			
	24				0·27	24	70			
20	0·5	3·44	150	125	2·74	125	250			
	1	2·46	107	124	1·98	90	248			
	2	1·74	76	121	1·41	64·5	242			
	4	1·26	55	115	1·05	48	230			
	8	0·96	42	105	0·85	39	210			
	16	0·74	32	86	0·68	31	172			
	24	0·60	26	70	0·55	25	140			
40	0·5	6·00	136	250						
	1	4·20	95	248						
	2	2·96	67	242						
	4	2·11	48	230						
	8	1·63	37	210						
	16	1·32	30	172						
	24	1·10	25	140						

0·025 μF. impulse generator.
Series resistance 2,000 ohms.
Air at atmosphere pressure.
Current decaying to half value in 28 μsec.

* Column (2) gives time from breakdown.

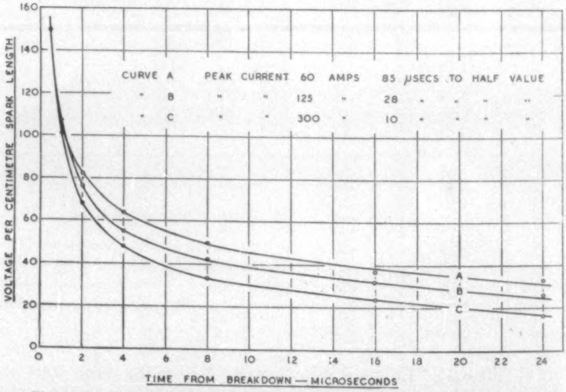


Figure 8. Voltage gradients in air for three different rates of current decay.

A few measurements were made for point–point and point–plane gaps with positive and negative polarity voltages applied. The results were practically identical and showed that the gap geometry does not affect the magnitude of the voltage drop. No differences were observed in the results obtained for electrodes of different materials, such as steel and brass, though such differences may be expected for measurements with short gaps when metal vapour is present in all parts of the discharge and not just in the region of the electrodes.

4.2. *Voltage Gradients in Nitrogen, Oxygen and Hydrogen at Atmospheric Pressure*

All the measurements were made for sparks across a 20 cm. gap between pointed electrodes, in a Perspex cylinder about 35 cm. in diameter and 90 cm. long. The gases used were of standard commercial purity.

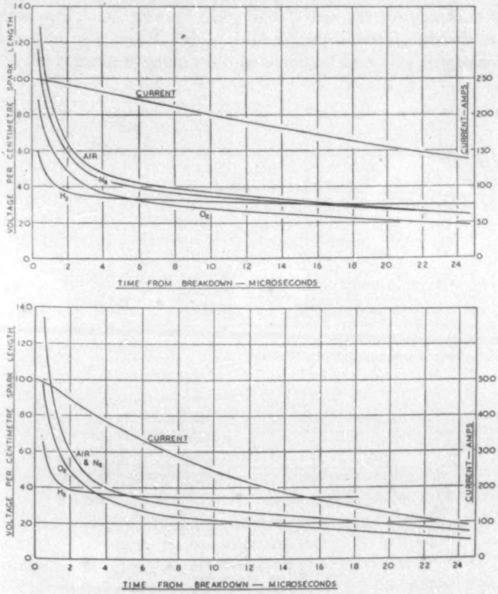


Figure 9. Voltage gradients in air, N_2, O_2, and H_2 with the associated current waveforms.

Figure 9 shows typical curves of voltage gradient as a function of time for nitrogen, oxygen, hydrogen and air. The characteristics for air, nitrogen and oxygen were found to be very similar in form, air always having the highest voltage gradient and oxygen the lowest, for peak values of current ranging between 125 and 500 amp. and for durations of 10 and 28 μsec. to half value. On the other hand, hydrogen always had an initially lower voltage gradient than those of the other gases but after two or three microsconds the voltage ceased to fall so rapidly and quickly reached a minimum. This is illustrated in the oscillograms of Figure 10

(Plate I) which show the recovery of voltage across the gap and the corresponding decay of the current to zero. Sparks in hydrogen also differed from those in the other gases in that the voltage gradient showed a dependence on the peak value of the current. For example, with currents decaying to half value in 10 μsec. the voltage gradients from 0·25 μsec. after breakdown to the time of commencement of voltage recovery were 40% greater for a peak current of 250 amp. than for a peak current of 500 amp.

4.3. *Voltages Gradients in Gases at Reduced Pressure*

A comprehensive series of measurements was made on sparks in air, nitrogen and oxygen at pressures from atmospheric down to 100 mm. Hg. Typical photographs of sparks in this pressure range are shown in Figure 3 which illustrates the general tendency for the spark channel to become straighter with reduced pressure (see also the graphs of Figure 4(*b*)). At 100 mm. Hg and particularly at lower pressures still, the spark is diffuse and the light emitted is relatively weak. Also the instant of breakdown is ill-defined due to the comparatively high current flowing during the formative stages of the spark (see oscillogram of Figure 11, Plate I) which produces a voltage drop in the series resistor. For this reason, 100 mm. Hg was the lower limit of pressure for which useful voltage measurements were made.

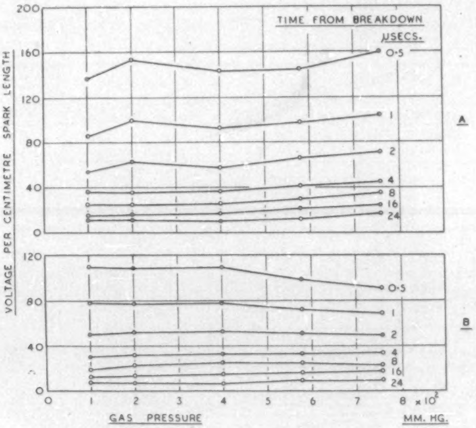


Figure 12. The effect of gas pressure on the voltage gradient. Current decays to half value in 10 μsec.
A. 20 cm. gap in air.
B. 20 cm. gap in oxygen.

The curves of Figure 12, in which voltage gradients at various times from breakdown are plotted against gas pressure, are typical of the results obtained, and show that variations of gas pressure, within the range investigated, has little effect on the voltage gradient.

4.4. *Sparks of Restricted Cross Section*

A series of voltage measurements was made on sparks in glass tubes and between glass plates.

Thick walled tubes, of Pyrex glass, with bores ranging between 1 and 6 mm. were used in the experiments. Sparks about 15 cm. in length were produced between steel wire electrodes, inserted loosely one at each end of the tube in use, and the voltage across the spark channel was measured. Typical voltage oscillograms are shown in Figure 13 (Plate I) and are characterized by an initial voltage decay, similar to that for an unrestricted spark, followed by a minimum and a subsequent rise in voltage. The initial region of voltage decay is not greatly affected by the presence of the glass tube, as indicated by the results of Figure 14.

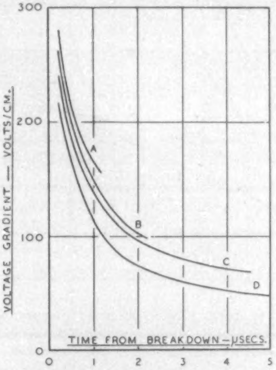


Figure 14. Initial voltages across spark channels restricted in glass tubes with the corresponding result in free air for comparison.

A. Tube bore 0·197 cm.
B. Tube bore 0·385 cm.
C. Tube bore 0·57 cm.
D. Free air.

It is reasonable to assume that the minimum in the voltage curve is related to the time when the spark channel has expanded to fill the tube, so that measurements with various tube bores should give an indication of the rate of expansion of the spark channel. The results for two currents are shown in Figure 15, curves A and B, the minima for the higher current being reached more quickly than for the lower current. The time was measured to the point where the voltage has decayed to within 5% of the minimum.

Voltage drops were also measured in sparks restricted between glass plates. It was found that the minimum occurred, but after a time greater than that required in the tubes. A result is shown in Figure 15, curve C. It is possible that the greater time arises because lateral expansion is possible in one direction after expansion has ceased in the other. As for glass tubes, a higher current produces a minimum in the voltage curve after a shorter time from breakdown.

The curves D and E of Figure 15 refer to spark channel diameter measurements on unrestricted spark channels made with the rotating mirror camera (Higham and Meek 1950) for the same two peak currents used in the glass tube experiments.

While the initial rate of expansion measured by the two methods is of the same order, the rate determined by the glass tube technique in the later stages is very much greater. According to the photographic results a spark channel conducting a peak current of 450 amp. reaches a maximum diameter of little more than 4 mm., whereas the tube measurements indicate that this diameter is reached in about 1·5 μsec. and that the expansion continues after this time at a nearly uniform rate. One possible explanation of the observed discrepancy between the results obtained by the two techniques is that the boundary of the visual photographed section of the channel and the boundary of the ionized region may not coincide. Curves A and B relate to the condition where the electrical characteristics of the spark become affected by the presence of the tube, and hence may be assumed to give the boundary of the ionized region, whereas curves A and E give the photographed boundary. According to this interpretation the effective diameter of the ionized region of a spark channel carrying 450 amp. at 2 μsec. after breakdown is 4·9 mm. while the effective diameter of the visible channel is only 2·4 mm. However, the presence of the tube may itself influence the expansion rate of the channel and further experimental study is required to confirm, or reject, the above proposals.

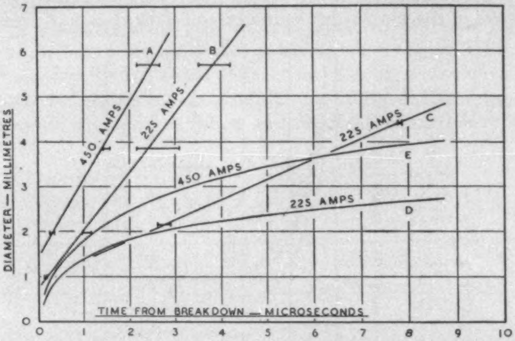


Figure 15. The expansion of spark channels in air as interpreted from voltage measurements on restricted channels. The results obtained photographically are included for comparison.

4.5. *Voltage Gradients in Spark Channels in Liquids*

A few measurements were made to determine the order of the voltage gradient for spark channels in liquids. Table 3 summarizes some results obtained for spark channels in ordinary tap water between brass pointed electrodes with a gap

Table 3. Voltage Gradients in volt/cm.

Peak current (amp.)	Time from breakdown (μsec.) 0.5	1	2	4	8	16	24
250	1700	1110	700	490	300	230	130
500	1280	920	630	420	300	240	160
750	1270	830	570	390	300	230	120

length of 2·5 cm. and currents decaying to half value in 10 μsec. No appreciable difference was recorded between the results for distilled and for ordinary tap water. The spark channel voltage gradients in transformer oil are of the same order as those in water.

The voltage gradients given in Table 3 denote the measured voltage drops divided by the length of the spark gap. It is probable that the sparks have a greater length than the gap, so that the true voltage gradients may be slightly lower than those given in the Table. If, as is the case for gaseous spark channels, the lengths of the sparks decrease with increasing over-voltage, the true voltage gradient for the currents investigated would be about 800 volt/cm. at 1 μsec. after breakdown and about 290 volt/cm. at 8 μsec. after breakdown. These values are about eight times those for air at atmospheric pressure.

§ 5. DISCUSSION

The present research concerning the voltage gradient along spark channels is one of a series of investigations into the spark channel now being performed in these laboratories. These investigations include measurements of the variation with time of the effective diameter of the spark channel and the radial distribution of light emission, and also spectroscopic studies from which ion densities may be determined. Until all the results of the various interrelated researches are known the interpretation of any single study is handicapped, particularly as there is little previous information available concerning the discharge processes involved. A complete theoretical treatment is therefore not attempted in this discussion which is intended primarily to draw attention to certain features of the results and to show some general correlation between them.

The change recorded in the length of the spark channel with variations in the magnitude of the applied voltage, for a fixed gap between the electrodes, is an effect which does not appear to have been noted previously. As the path taken by the leader stroke determines the shape of the spark channel (Allibone and Meek 1938), the changes in channel length can be explained only by consideration of the leader stroke process. No complete theoretical analysis of the mechanism of the leader stroke is yet available (see, however, Schonland 1938, Meek 1939, Bruce 1944, Loeb 1948), but in general it is assumed that the leader stroke proceeds as a consequence of electron avalanche development in the region of its tip. As the electron avalanches are initiated in a random manner, and can grow within a large solid angle surrounding the tip of the leader stroke, the leader stroke develops along a zig-zag path.

If the voltage applied to the gap is increased above the minimum value required to cause breakdown, the number and size of electron avalanches produced in the region surrounding the tip may be expected to increase and there is an enhanced probability of successful avalanche development along the line of greatest field strength between the tip of the leader stroke and the opposite electrode.

With reduction in gas pressure, the mean free path of electrons is increased, so that the electron avalanches are longer, and as the leader stroke is made up of a succession of avalanches it may itself be expected to take a less irregular course. Also, at lower gas pressures there is greater diffusion during the growth of avalanches and streamers or leader strokes, and the path of the leader stroke assumes a smoother and more rounded form. The more direct path taken at reduced pressures would also seem to imply an increase in the number of avalanches produced round the tip of the leader stroke, so that there is a greater chance of successful avalanche growth along the line between the tip and the opposite electrode.

An increase in spark length at pressures higher than atmospheric has been recorded by Howell (1939) for breakdown between a point and a plane. In his experiments an A.C. voltage was used and considerable space charges were produced before breakdown occurred, with consequent field distortion which may have affected the leader stroke development. In the present experiments, the time between the application of the impulse voltage and the breakdown of the gap is too short for drift of space charge across the gap.

All the present measurements have been made for point–point gaps, for which simultaneous growth occurs of leader strokes from both positive and negative points (Allibone and Meek 1938). Differences may be expected for point–point gaps, with points of positive or negative polarity, or for sphere–sphere gaps, and further relevant studies are being undertaken.

The voltage drop varies directly with the gap length and hence the voltage gradient along the spark channel may be assumed to be nearly uniform. There is no reason to suppose that one portion of the spark channel should differ from another, except in the region of the electrodes, or where the direction of the spark path is sharply inclined to the line joining the two electrodes. For instance, the gradient at bends of the channel may differ from that of adjacent portions of the channel, but on balance these variations probably cancel out and give a resultant average voltage gradient for the whole channel varying little from spark to spark.

The magnitudes of the voltage gradients recorded in the various gases are appreciably higher than those for arcs carrying the same current. Measurements by von Engel (1929) for an arc in air, 10 cm. in length and carrying a current of several hundred amperes, show that the voltage gradient is about 2 volt/cm. The present results for sparks carrying a current of approximately the same value give a voltage gradient of about 100 volts/cm. at a time of 1 μsec. after breakdown and about 40 volt/cm. at a time of 8 μsec.

It is clear that a higher voltage gradient would be expected in a spark than in an arc, as in the former case equilibrium has not been established. In a spark channel the gas density is initially the same as that of the surrounding gas, but with the sudden input of energy the effective gas temperature rises rapidly, to a value thought to be of the order of 10,000° C. (Craggs and Meek 1946), with a consequent increase in gas pressure and a rapid expansion of the channel. This expansion continues, and conditions within the channel change, until the gas pressure within the channel falls to that of the surrounding gas. Meanwhile, if the current is maintained, the gas density falls and the temperature assumes a value of about 6,000° C. for arcs in air at atmospheric pressure. On the above basis it is possible, in some respects, to liken the spark to the case of an arc which is artificially constricted to a cross section appreciably smaller than the normal value.

In the light of the above general picture of the spark it is interesting to draw attention to some of the trends observed in the present measurements. One distinctive feature is that over a wide range of currents there is little change in the value of the voltage gradient at given times from breakdown. For instance, as shown in Table 1, for a gap of 10 cm. in air, the voltage gradient in channels carrying a peak current ranging from 150 to 450 amp., and decaying to half value in 10 μsec.

varies between 142 and 165 volt/cm. at 0·5 μsec. after breakdown, and between 34·5 and 36·5 volt/cm. at 8 μsec. after breakdown. Measurements of spark channel diameters (Higham and Meek 1950) indicate that at 8 μsec. after breakdown the corresponding sparks, carrying peak currents of 150 and 450 amp., have diameters of 0·21 and 0·39 cm. respectively. The average current densities in these two cases are then 2·5 and 2·2 ka/cm^2 and are hence roughly equal. From analysis of these and other records of spark channels in air, it appears to be true, to a first approximation, that at a given time after breakdown the voltage gradients for sparks carrying different currents are the same and, further, that the current densities at a given time are the same. Consequently, for a given current wave shape, the conductivity of a spark channel is roughly constant at a given time from breakdown for a wide range of peak currents, between 100 and 1,000 amp. approximately. For currents rising to peak value in about 0·25 μsec., and decaying to half value in times longer than 10 μsec. the value obtained for the conductivity is about 150 ohm^{-1} cm^{-1} at a time of 1 μsec. from breakdown. At a time of 8 μsec. from breakdown the value is more dependent on current wave shape and values of 75 ohm^{-1} cm^{-1} and 90 ohm^{-1} cm^{-1} were obtained for currents decaying to half value in 10 μsec. and 28 μsec. respectively. Norinder and Karsten (1949) quote a conductivity of 400 ohm^{-1} cm^{-1} for sparks carrying oscillatory currents of peak amplitude between 10 and 100 kiloamp.

The product of current density and voltage gradient gives the power input per unit volume of the spark channel. For the spark channels in air the power input at 1 μsec. after breakdown is of the order of 1 megawatt per cm^3.

The close connection between voltage gradient and channel diameter is also shown by the results for gases other than air. The voltage gradients recorded in nitrogen are closely the same as for air and the channel diameter expands at a similar rate (Higham and Meek 1950). In oxygen the voltage gradient is slightly lower than in air but the channel diameter is greater, and consequently the current density is lower. In hydrogen the voltage gradient is lower than in air for the early stages of the spark channel, and is higher in the later stages; corresponding observations of the rate of expansion of the channel (Higham and Meek 1950) show that the expansion is more rapid in the early stages and slower in the later stages.

The results obtained for the sparks in tubes show that the voltage gradient is increased by the artificial constriction of the channel cross section. Even before the spark channel has expanded to fill the tube, the walls may be expected to influence the characteristics of the spark channel appreciably, partly because of their effect in the rate of energy loss but also because of the reflections of the pressure wave generated when the spark is initiated. For such reasons it is unwise to rely too greatly on any deductions concerning the rate of growth of channel diameters from a comparison between the voltage drops for sparks in free air and for sparks in tubes. The comparatively high gradients in spark channels in liquids are probably caused largely by the constriction of the channel, which cannot expand freely because of the opposition from the liquid boundary.

The voltage gradient measurements of Figure 8, which show the effect of changing the rate of current decay, have been extrapolated to determine the voltage gradients for a unit function current, i.e. a current rising to its peak magnitude instantaneously and thereafter remaining constant. The results are given in

Table 4, and may be expected to apply with reasonable accuracy to spark channels in air at atmospheric pressure when conducting unit function currents of magnitudes ranging between 50 to 1,000 amp.

Table 4

Time from breakdown (μsec.)	0.5	1	2	4	8	16	24
Voltage gradient (volt/cm.)	150	118	93	72	56	43	35

In future work it is intended to overcome the complication of current decay by the use of square current pulses which rise to a peak in a time of the order of 0·1 μsec. and remain at that value for a prescribed period, of possibly 10 μsec. or more. Experiments with square pulses have been made by various investigators, but in all cases the voltage available was such that it was not possible to obtain breakdown of gaps of more than about 2 cm. With the impulse generator used in the present experiments, the voltage available, of 1 million volts, is sufficient to produce breakdown gaps up to about 150 cm. but it is not possible to produce a square current pulse of amplitude 1,000 amp. or more in the spark channel. A new impulse generator is now being built with which it is hoped to produce pulses of 100 kv. giving a current of, for example, 1,000 amp. lasting for 10 μsec.

A further improvement in technique may also be adopted in future investigations in order to determine more accurately the voltage drop in the early stages of the channel, within 0·2 μsec. from breakdown. The actual instant of breakdown is not readily recorded to within less than 0·1 μsec. because of the character of the oscillogram, and consequently the measurement of the voltage drops in the early stages of spark formation is inaccurate. One possible method of improving the records is to use two synchronized oscillographs, one within a linear divider and the other with a non-linear divider. From the former a more precise record of the breakdown time can be obtained, while from the latter the voltage drop may be measured.

REFERENCES

Allibone, T. E., and Meek, J. M., 1938, *Proc. Roy. Soc.* A, **166**, 97, **169**, 246.
Bruce, C. E. R., 1944, *Proc. Roy. Soc.* A, **183**, 228.
Craggs, J. D., and Meek, J. M., 1946, *Proc. Roy. Soc.* A, **186**, 241.
von Engel, A., 1929, Z, *Tech. Phys.*, **10**, 505.
Flowers, J. W., 1943, *Phys. Rev.*, **64**, 225.
Higham, J. B., and Meek, J. M., 1950, *Proc. Phys. Soc.* B, **63**, 649.
Howell, A. H., 1939, *Trans. Amer, Inst. Elect. Engrs.*, **58**, 193.
Jones, F. Llewellyn, 1945, *J. Soc. Chem. Ind.*, **64**, 317; 1946, *Nature, Lond.*, **157**, 298.
Loeb, L. B., 1948, *J. Franklin Inst.*, **246**, 123.
McCann, G. D., and Clark, J. J., 1943, *Trans. Amer. Inst. Elect. Engrs.*, **62**, 45.
Meek, J. M., 1939, *Phys. Rev.*, **55**, 972; 1947, *Nature, Lond.*, **160**, 110.
Norinder, H., and Karsten, O., 1949, *Ark. Mat. Astr. Fys.*, **36A**, Part 4.
Prime, H. A., and Saxe, R. F., 1949, *Proc. Instn. Elect. Engrs.*, **96**, Pt. II, 662.
Schonland, B. F. J., 1938, *Proc. Roy. Soc.* A, **118**, 233.
Williams, G. C., Craggs, J. D., and Hopwood, W., 1949, *Proc. Phys. Soc.* B, **62**, 49.

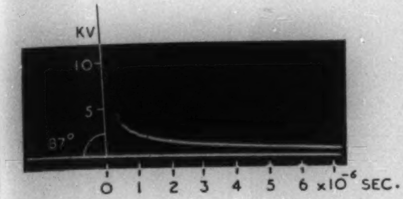


Figure 2. Oscillogram of voltage drop across spark channel on which the Y axis has been drawn. (The original negative shows a clearly defined trace above 10 KV.)

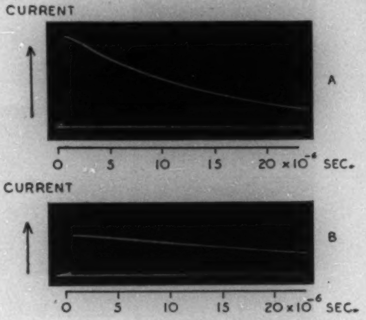


Figure 7. Typical current oscillograms.
A. Current decaying to half value in 10 μsec. with 750 ohms series resistance.
B. Current decaying to half value in 28 μsec. with 2,000 ohms series resistance.

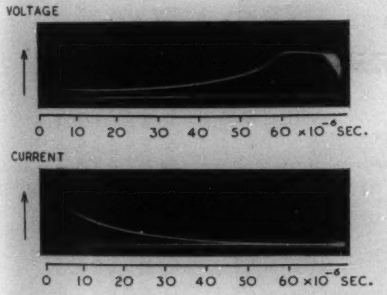


Figure 10. Oscillograms showing the recovery of voltage across a hydrogen spark channel with the corresponding decay of current to zero. After current extinction the voltage decays by current flowing through the potential divider in parallel with the spark gap.

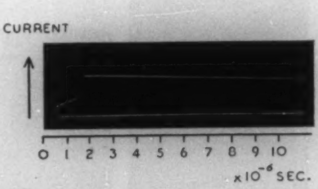


Figure 11. Oscillogram showing the current flowing through a spark channel in air at 200 mm. Hg. (An appreciable current flows before complete breakdown occurs at $1{\cdot}8\times10^{-6}$ sec.)

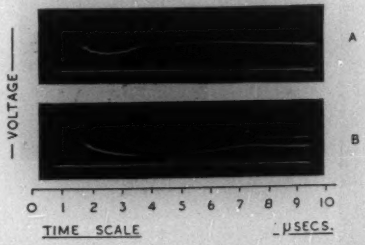
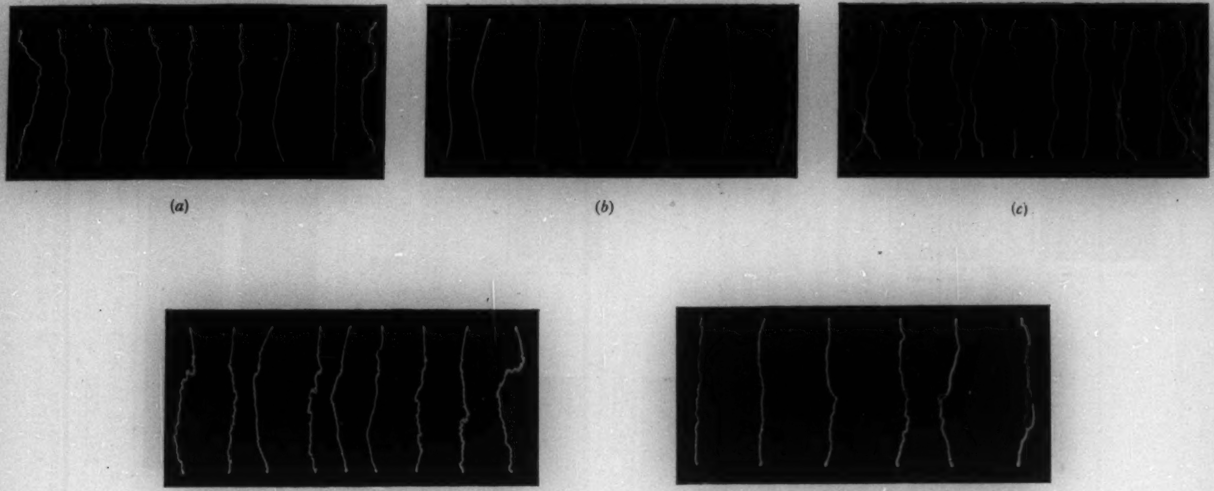


Figure 13. Oscillograms showing the voltages across spark channels restricted in glass tubes. A. 0·197 bore. B. 0·385 cm. bore.

PLATE I.



(*a*) (*b*) (*c*)

(*d*) (*e*)

Figure 3. Typical photographs of sparks in 20 cm. gaps.

(*a*) Nitrogen at atmos. pressure. Applied voltage 200 kv. Peak current 250 amp.
(*b*) Nitrogen at 200 mm. Hg. Applied voltage 200 kv. Peak current 250 amp.
(*c*) Air at atmos. pressure. Applied voltage 200 kv. Peak current 250 amp.
(*d*) Hydrogen at atmos. pressure. Applied voltage 200 kv. Peak current 250 amp.
(*e*) Oxygen at atmos. pressure. Applied voltage 490 kv. Peak current 500 amp.

(The second images in these photographs are caused by reflection from the wall of the cylindrical Perspex gas chamber.)

PLATE II.

PROC. PHYS. SOC. VOL. 63, PT. 9—B (J. B. HIGHAM AND J. M. MEEK)
'The Expansion of Gaseous Spark Channels'

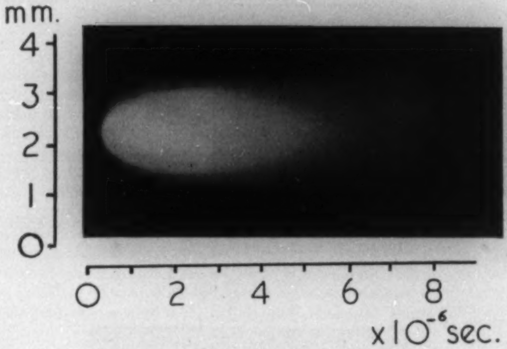


Figure 2. Rotating mirror photograph showing expansion of spark channel.

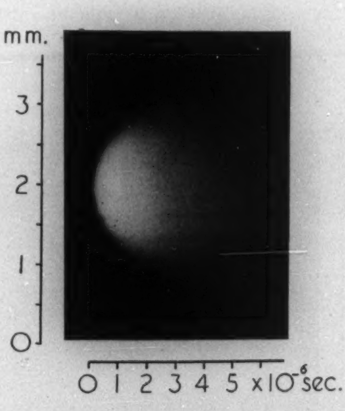


Figure 3. High resolution rotating mirror photograph showing initial expansion of spark channel.

PLATE I.

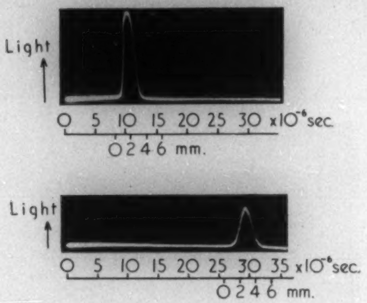


Figure 5. Typical oscillograms showing the distribution of light across spark channel.

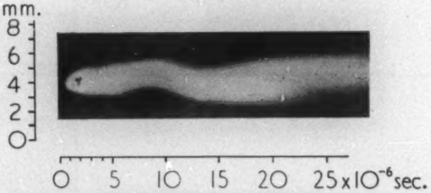

mm.
8
6
4
2
0
0 5 10 15 20 25 x10⁻⁶sec.

Figure 12. Rotating mirror camera record showing a typical hydrogen spark channel expansion.

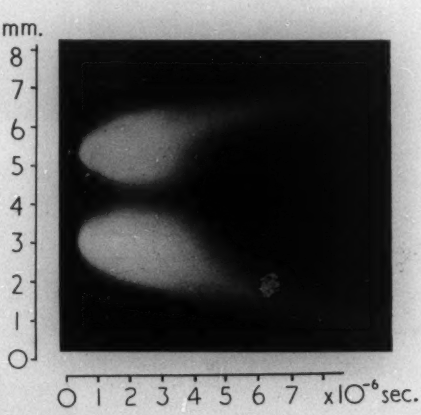

mm.
8
7
6
5
4
3
2
1
0
0 1 2 3 4 5 6 7 x10⁻⁶sec.

Figure 14. Rotating mirror camera photograph of a branched spark channel in nitrogen conducting a current of 500 amps.

The Expansion of Gaseous Spark Channels

By J. B. HIGHAM and J. M. MEEK
Department of Electrical Engineering, University of Liverpool

MS. received 9th March 1950

ABSTRACT. A rotating mirror camera with a temporal resolution on the film of 1·5 mm/μsec. has been used to determine the rate of expansion of long gaseous spark channels during their first 10 μsec. of growth. The spark channels conducted aperiodic impulse currents in the range of 60 to 500 amp. peak, which attained their peak values in about 1/4 μsec. and decayed to half-value in 10 or 28 μsec. Sparks have been investigated in air at atmospheric and reduced pressures, and in nitrogen, oxygen and hydrogen at atmospheric pressure. The measured areas of cross section of these spark channels bear a linear relation to the peak currents, except in hydrogen, and are little influenced by the rate of current decay in the range studied. Measurements have also been made with the rotating mirror scanning an image of the spark across a photomultiplier connected to a high speed oscillograph; by this method a direct record is obtained of the radial light distribution across the spark channels. The mechanism of spark channel expansion is discussed, and estimates are made of the average ion densities in spark channels.

---

§ 1. INTRODUCTION

Until the studies by Flowers (1943), who recorded the luminous boundary of the spark channel by a rotating-film camera, there have been few investigations of the current densities occurring in spark channels, though a certain amount of information is available from earlier experiments by Beams (1930), Lawrence and Dunnington (1930) and Dunnington (1931), all of whom used a Kerr cell type of electro-optical shutter. Other measurements with rotating cameras have been made by Raiskij (1940, 1948), and more recently by Norinder and Karsten (1949).

In the present investigation (Meek 1947), a similar technique to that employed by Flowers (1943) has been adopted, but an improvement in time resolution has been obtained by the use of a rotating mirror instead of a moving film. This has enabled records to be made of the spark channel at earlier times subsequent to the formation of the channel.

As in the previous studies of spark channels it is assumed in the analysis of the results that the luminous boundary of the spark channel defines also the boundary of the region of ionization and therefore of the current flowing. This assumption has not been justified experimentally, but it is frequently adopted in electrical discharge measurements, and Suits (1935) considers that it is justifiable in the case of steady high-pressure arcs. Indirect evidence showing that the electrical boundary is greater than the visible boundary is given in the measurements of voltage gradient along sparks in tubes, but such evidence cannot be considered conclusive (Higham and Meek 1950). In the present records, although the light intensity varies across the spark channel, the luminous boundary is quite sharply defined and enables an effective diameter to be assigned to the spark. While the current density will also vary radially over the channel, an average effective current density may be defined as the total current divided by the effective cross-sectional area of the spark, and this definition has been used in the analysis of the results.

§ 2. MEASUREMENT TECHNIQUES

The present work was undertaken in parallel with the investigations on the voltage gradients in long spark channels (Higham and Meek 1950), and consequently the sparks used were of the same nature in the two cases. The point–point spark gaps were supplied from an 8-stage impulse generator of discharge capacitance 0·025 μF. through series resistors of 750 or 2,000 ohms. The currents attained their peak values in about 0·25 μsec., decayed to half-value in 10 or 28 μsec., and their peak amplitudes ranged from 60 to 500 amp.

2.1. *Rotating-Mirror Camera*

The optical arrangement of the rotating-mirror camera used for the measurements is shown diagrammatically in Figure 1. An image of the spark is formed by an object lens on an adjustable slit, of the type used in spectrographic work. The light passing through the slit from a short length of the spark is collected by a second lens and, after reflection by a rotating mirror, is brought to a focus on the photographic film. The resultant record then shows the photographed spark channel diameter as a function of time. The overall magnification from spark to film usually employed was 2 to 3 diameters, and, for a mirror rotational speed of 12,000 r.p.m., the time scale on the film is about 1·5 mm/μsec., about four times faster than that used by Norinder and Karsten (1949). The rotational speed of the mirror was measured by a stroboscope.

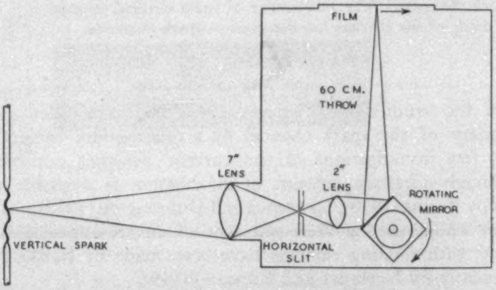


Figure 1. Schematic diagram of the rotating mirror camera.

A typical record showing the expansion of a spark channel is reproduced in Figure 2 (Plate I). For the purpose of analysis, the records were enlarged by about 10 diameters and tracings were prepared.

A necessary feature of the equipment is the arrangement for synchronizing the mirror position and the initiation of the spark so that when the spark occurs the mirror is correctly placed to reflect an image of the spark on to the film. For this purpose, a collimated light beam is reflected once each revolution of the mirror on to a photoelectric multiplier which is normally biased off. When the charging voltage of the impulse generator reaches a pre-determined value, the bias is automatically removed and the resultant pulse from the multiplier is amplified and fed to a blocking oscillator which provides a triggering spark (Husbands and Higham 1950) for the impulse generator. The instantaneous position of the mirror when the spark occurs can then be controlled so that a record is obtained from each spark. With an 8-stage impulse generator the occurrence of the spark was controlled to within about 3 μsec., represented by only 5 mm. on the 20 cm. length of exposed film. Full details of the circuits involved are published elsewhere (Prime and Saxe 1949).

Trial records were made with slits of various widths such as to give image widths in the plane of the film ranging from 5 mm. down to 0·3 mm., corresponding to exposure durations of 3·3 down to 0·2 μsec. It was found that the channel diameters recorded with this range of slit widths were identical during the first few microseconds. The use of a narrow slit serves to improve the definition of the record for the initial stages of growth, as illustrated in Figure 3 (Plate I) which shows a record obtained with a slit image of 0·5 mm. Because of the short exposure involved, the channel is not recorded in its later stages, and it was found necessary to use a larger slit width. The slit usually employed gave an image in the plane of the film 2·5 mm. wide, corresponding to an exposure of 1·7 μsec., and was found suitable for most of the sparks investigated. The record of Figure 2 was obtained with this slit. All the negatives show a clearly defined boundary to the light-emitting channel.

Both lenses in the camera were operated at f/2·8, and the 35 mm. film employed was Ilford H.P.3 fast panchromatic which was developed in MQ developer under standardized time and temperature conditions. This standardization enabled reasonable comparisons to be made between the photographed diameters of sparks for different currents and gas pressures. No attempt was made to compare actual light intensities, though this will be done in a future investigation after further work has been done in calibrating the film, as little is known on the response of photographic emulsions to the short duration exposures involved.

Analyses were made of about eight sparks for each set of conditions investigated, and it was found that the results were closely consistent throughout. The variation in measured diameters was rarely more than ± 0·15 mm. from one record to another.

2.2. *Photomultiplier–Oscillograph Recording Equipment*

A technique developed in the later stages of the research programme overcomes some of the difficulties inherent in photographic measurements. Some details of this technique have already been published (Saxe and Higham 1950). The rotating mirror camera is used without its resolving slit to scan an image of the spark across a slit in front of a photomultiplier, the spark and the photomultiplier slit being parallel to the axis of rotation of the mirror. The arrangement is shown schematically in Figure 4. By connecting the photomultiplier output through an amplifier (Prime and Saxe 1949) to a high speed oscillograph a record may thus be obtained of the light emission across a spark channel. Typical oscillograms are reproduced in Figure 5 (Plate II).

As in the photographic technique, the occurrence of the spark is synchronized with the mirror position, so that the image of the spark can be made to scan across the slit at any time relative to the initiation of the spark.

This technique, which was developed primarily for a future investigation to determine the distribution of light across a spark channel, also provides a check on the channel diameters as determined photographically in the present work.

A typical set of results is plotted in Figure 6 with the corresponding photographic results for comparison.

The photomultiplier method is more adaptable than the photographic method for investigating spark channels in their later stages when the light emission is weak, though photographic results can be obtained for these later stages by using a wider resolving slit. However, the method is not so readily applicable to the initial stages of channel growth, when the diameter changes appreciably during the time of scan. A serious drawback to the multiplier technique is that it provides no definite means of indicating whether the spark is being scanned perpendicularly or obliquely to its axis. It was possible to see in the photographs whether the spark was perpendicular to the resolving slit, and if it was not perpendicular the record was discarded. This is probably the major factor causing most of the diameters determined by this method to be greater than the photographed diameters as shown in Figure 6. A further advantage of the photographic technique is that it provides a continuous record of diameter as a function of time, whereas only one cross section at a given time from breakdown is provided by each multiplier oscillogram.

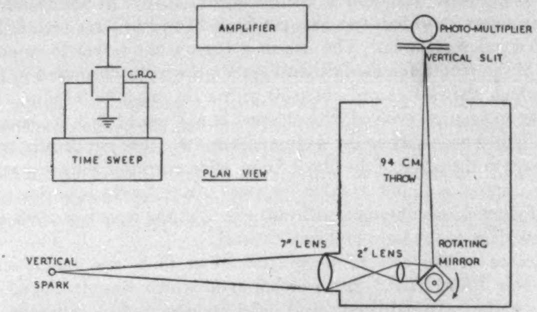


Figure 4. Schematic arrangement of the rotating mirror–photomultiplier–oscillograph recording equipment.

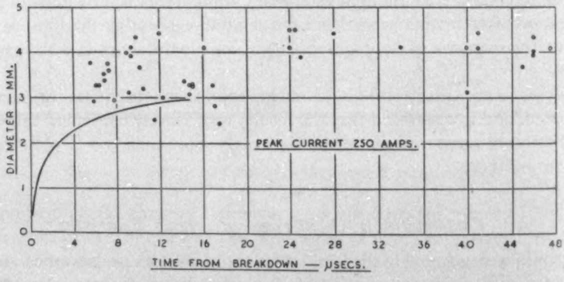


Figure 6. Spark channel diameters in air obtained with the rotating mirror–photomultiplier technique. The full curve represents results obtained with the photographic technique.

§ 3. RESULTS

The results presented in §§ 3.1 to 3.3 refer to the spark channel diameter as photographed with the rotating mirror camera, though reference is made in § 3.1 to some results obtained by the photomultiplier technique.

3.1. *Air at Atmospheric Pressure*

Preliminary measurements made on sparks of various lengths from 1 to 40 cm. showed that the length had no influence on the channel diameter. Further, the diameter appears constant along the length of the channel except at positions close to the electrodes where the channel is affected by the presence of electrode vapour. Consequently it was decided to standardize on one gap length, of 10 cm., and to make measurements at a point in the channel about 1 cm. from one of the electrodes.

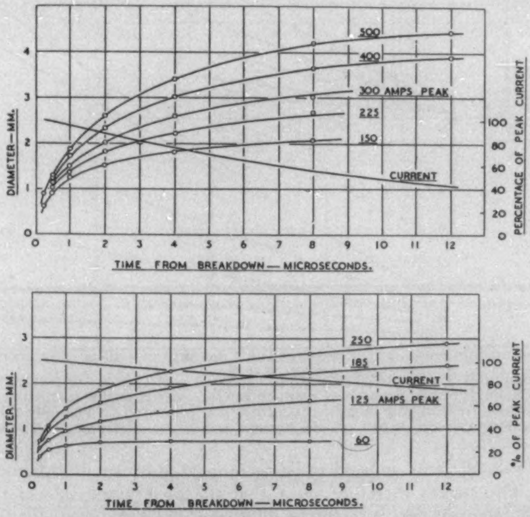


Figure 7. Photographically determined spark channel diameters in air as a function of time for various values of peak current and two rates of current decay.

A series of records was obtained for sparks across the 10 cm. gap for two current waveforms and for various values of peak current. The measured channel diameters are shown as a function of time in the two graphs of Figure 7 together with the corresponding current waveforms. The initial growth of the spark channel is very rapid, and, for a spark conducting a peak current of 500 amp. the diameter is increasing at a speed of about $9{\cdot}3 \times 10^4$ cm. per sec. at 1 μsec. after the formation of the channel. At 8 μsec. from breakdown the rate of expansion has decreased to about $1{\cdot}2 \times 10^4$ cm. per sec. With lower currents in the spark the initial rate of expansion is less rapid, being about $4{\cdot}6 \times 10^4$ cm. per sec. at 1 μsec. after breakdown for a current of 150 amp.

However, after 10 μsec. the rates of expansion for the sparks carrying peak currents ranging between 150 amp. and 500 amp. are closely the same.

In Figure 8 the measured cross-sectional areas are plotted as a function of peak current, with time from breakdown as the parameter. For each rate of current decay employed the cross section is found to increase in a nearly linear manner with the peak current value. This implies that the current density in the sparks observed for a given rate of current decay is approximately constant at a given time from breakdown. With a peak current of 250 amp. the channel dimensions measured for the two rates of current decay are almost identical.

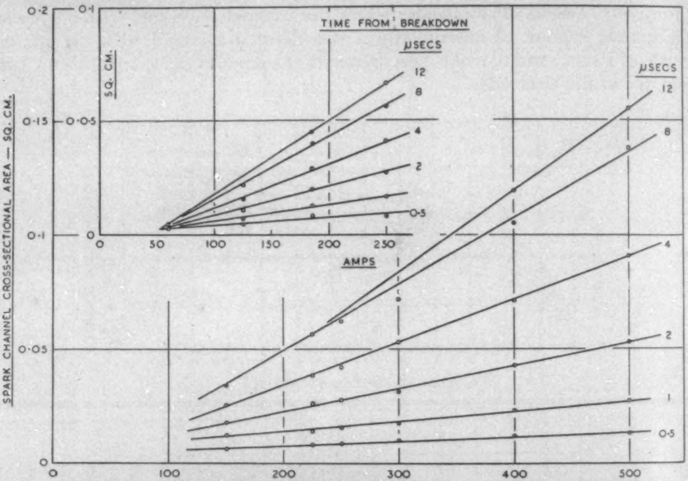


Figure 8. Photographically determined spark channel cross-sectional areas plotted as a function of peak current, with time from breakdown as parameter. The upper graph refers to currents decaying to half value in 28 μsec., and the lower graph to half value in 13 μsec.

The lower rate of decay appears to be accompanied by smaller diameters below 250 amp. The curves therefore show that the peak amplitude of the current is a decisive factor in determining the diameter, while the rates of current decay for the two cases investigated are less important.

Some values determined for the average current densities in sparks in air at various times from breakdown are given in Table 1.

Table 1

Time from breakdown (μsec.)	0·5	1	2	4	8
Current	250 amp. peak decaying to half-value in 10 μsec.				
Current density (k.amp/cm^2)	28·6	14·5	7·8	4·5	2·4
Current	250 amp. peak decaying to half-value in 28 μsec.				
Current density (k.amp/cm^2)	29·0	15·0	8·9	5·6	3·7

The current densities in all the sparks recorded in air at a given time from breakdown are of the same order of magnitude as those given in Table 1, though higher densities were observed in the later stages of a channel conducting a peak current of 60 amp., decaying to half-value in 28 μsec. This was the lowest peak current used in the experiments, and it is possible that the photographed image of the channel gives too small a diameter, because of the comparatively weak light emission, with a consequent apparent increase in the current density.

While the purpose of the investigations is primarily to determine the effective diameter of the spark channel, some measurements were also made, with the photomultiplier technique, of the variation of the light emission across spark channels in air. A number of oscillograms were analysed and the results are plotted in Figure 9 where the the ordinate gives relative light intensity only, with unit intensity at the centre of the channel. The oscillograms correspond to times between 10 and 30 μsec. after breakdown for spark channels carrying currents from 60 to 250 amp. It is clear that the distribution of light is closely the same for the range of conditions investigated.

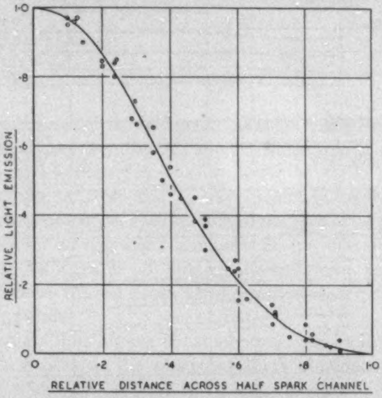


Figure 9. Results (to arbitrary scale) for the variation in light emission across a number of spark channels in air conducting currents ranging between 60 and 250 amps and at times from breakcown from 10 to 30 μsec.

3.2. *Air at Reduced Pressure*

Some of the measured spark channel diameters for air at reduced pressures are shown in Figure 10. It will be seen that down to 400 mm. Hg there is little change in the photographed diameter. Visual observation shows, however, that the channels produced at 100 and 200 mm. Hg are diffuse in nature, and this is borne out by the increased diameter recorded for the channel carrying a peak current of 500 amp. For channels carrying less current than this the light emitted is so weak that the records obtained are ill-defined for the later stages of growth, and these results have consequently been omitted from the graphs.

3.3. *Nitrogen, Oxygen and Hydrogen at Atmospheric Pressure*

Experiments were made with the 10 cm. spark gap enclosed in a Perspex cylinder 35 cm. in diameter and 90 cm. in length. Gases of commercial purity were used.

The results obtained for the various gases are shown graphically in Figure 11 with the corresponding results for air included for comparison. As for air, the important factor determining the channel diameter in the other gases studied is the peak current and not the rate of current decay. In nitrogen and oxygen the spark channels expand at about the same rate as that observed for air, and therefore the current densities in sparks in these three gases are closely similar.

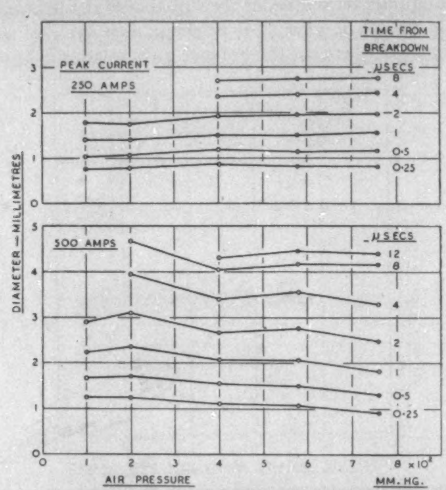


Figure 10. The effect of gas pressure on photographically determined spark channel diameters in air.

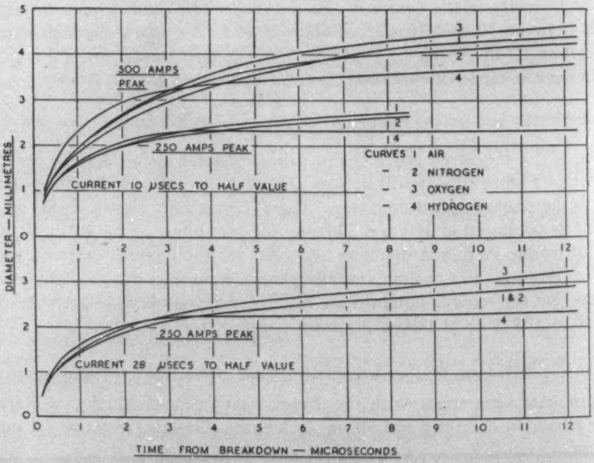


Figure 11. Photographed spark channel diameters in air, N_2, O_2 and H_2.

In hydrogen the expansion is more rapid in the early stages, but after about 5 μsec. the channel reaches a nearly stable diameter. The result is that the current density is lower in hydrogen than in air during the first few microseconds after breakdown, but is higher in the later stages. Also, in contrast with air, the average current density in hydrogen appears to be greater for lower peak currents in the spark than for higher peak currents, as shown by Table 2, which gives the voltage gradients and current densities for sparks in hydrogen for currents decaying to half-values in 10 μsec.

Table 2

Peak current (amp.)	250		500	
Time from breakdown (μsec.)	2	8	2	8
Voltage gradient (v/cm.)	55	46·5	33·5	34·5
Current density (k.amp/cm^2)	6·95	3·6	6·55	2·9

Three other features of the hydrogen channel differentiating it from the channels in the other gases were noted, and are illustrated to some extent by the photograph of Figure 12 (Plate II).

(i) The light emission from the spark channel in hydrogen is practically constant until the current has decayed to a small fraction of its peak value, whereas for a spark in air under similar conditions of current, camera slit width etc., as illustrated by the photograph of Figure 2, the light emission has fallen considerably after about 10 μsec. from breakdown when the current has decayed to half its peak amplitude.

(ii) The light emission is almost constant across the hydrogen spark channel, but attains a definite peak at the centre of spark channels in the other gases investigated. This is clearly shown by the photographic density measurements of Figure 13 which refer to typical spark channels in air and hydrogen at about 3 μsec. after breakdown.

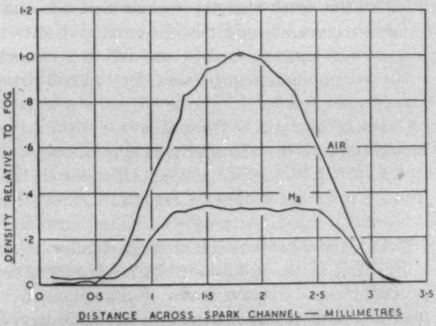


Figure 13. Photographic density measurements across rotating-mirror camera photographs of spark channels in air and hydrogen 3 μsec. after breakdown. The current was the same in each case.

(iii) The photographs show that the long spark channels in hydrogen oscillate laterally. The amplitude of oscillation is approximately equal to the channel radius, and the frequency is of the order of 100 kc/s. Most of the records exhibited this phenomenon but it is not possible to say definitely from the present results whether the oscillation is in one plane or is a circular motion around the average axis of the spark. The latter seems the more likely, and would account for so many of the records showing the same amplitude of oscillation.

A rotating mirror camera photograph showing the expansion of a branched spark channel in nitrogen is given in Figure 14 (Plate II). 30% of the nitrogen records exhibited this effect, but it was rarely observed in the other gases. Branching on a larger scale, where the branch channels are well separated, was observed no more frequently in nitrogen than in the other gases.

A point of particular interest about these closely adjacent branches is the manner in which they expand, as all the records obtained show that the brightest portions of the channels decrease in cross section and move away from each other. Close examination of the original negatives reveals a dark layer of nearly constant thickness between the two branches. In the record shown there also appear two streamers of short duration close to the main discharge in its early stages of growth.

§ 4. DISCUSSION

As explained in the preceding paper (Higham and Meek 1950), a spark channel is initiated in a discharge gap when the space between the two electrodes has been bridged by a conducting plasma known as the leader stroke. The external circuit then commences to discharge through this plasma, producing the so-called main stroke (Schonland 1938, Allibone and Meek 1938). The current flowing in the main stroke depends on the external circuit but is generally many times higher than that in the leader stroke. The initial sectional area of the main stroke may then be assumed to be that of the leader stroke, but, because of the greater current and power expended in the main stroke than in the leader stroke, there is a rapid increase in the ionization within the channel with a consequent rise in temperature. The energy supplied to the spark channel is expended not only in producing the necessary ionization required to maintain the current in the channel but also in compensating the various losses caused by the diffusion of electrons and ions from the channel, by recombination processes, by light radiation and by the sound wave produced by the spark.

No complete theoretical analysis of the conditions within the spark channel is yet available, though discussions of various aspects of the problem are given in papers by Toepler (1925, 1926, 1927, 1929), Ollendorff (1933), Fucks and Bongartz (1943), Rompe and Weizel (1944), Craggs and Meek (1946), Craggs and Hopwood (1947), Norinder and Karsten (1949). The most detailed treatment is that by Ollendorff (1933) who has considered the particular case of the lightning channel, and, on the basis of the Saha equation and assumptions concerning diffusion, has deduced certain characteristics of the channel. However, the calculations are made for a discharge assumed to be stationary with respect to time, and no consideration is given to the rate of expansion of the channel.

Among the various factors governing the growth of the spark channel it is probable that both diffusion and recombination play important rôles, diffusion principally in determining the rate of expansion and recombination in influencing the density of ionization within the channel. Hopwood (unpublished work) has suggested that the liberation of photons from within the channel may also cause photo-ionization of the gas surrounding the channel and so extend the region of ionization, in this way causing an effective expansion of the channel.

In the initial stages of the spark the temperature of the electrons will be appreciably higher than that of the neutral atoms and molecules, but, as an electron makes approximately 10^5 collisions per μsec. with molecules in air at atmospheric pressure, the gas temperature, and therefore the gas pressure, within the channel will be greatly increased within a time of the order of 1 μsec. Diffusion of electrons and positive ions will therefore occur from the ionized region, at a rate depending on the magnitude of the diffusion coefficients applicable to the particular conditions concerned. While electrons tend to diffuse at a more rapid rate than the positive ions, because of their higher speeds, the resultant excess positive space-charge field produced by positive ions left behind produces a field tending to retard the escaping electrons, and hence the expansion rate is probably governed by the so-called 'ambipolar' diffusion coefficient rather than by the diffusion coefficient for electrons alone. This expansion rate clearly relates to the electrical boundary of the channel and, as mentioned in § 1, may or may not coincide with the recorded visible boundary.

Although it is probable that diffusion is the principal factor causing the radial expansion of the spark channel, calculations on this basis alone are not yet possible because of the present lack of knowledge of the distribution of ionization in the spark channel, and are further complicated by the fact that the supply of energy to the spark channel is continually changing. The diffusion rate may also be influenced by the high magnetic field surrounding the spark channel ; for instance, the magnetic field strength at the circumference of a channel of diameter 1 mm. when carrying a current of 1,000 amp. is 4,000 oersteds. However, it has been pointed out by Craggs that the problem is greatly simplified in the case of the expansion of channels during the afterglow period immediately following the cessation of current flow, and calculations on this basis are being made in conjunction with experiments on sparks in hydrogen by Craig and Craggs (1950).

As the diffusion coefficients for positive ions in air, oxygen and nitrogen are roughly the same, the rates of expansion of spark channels in these gases may then be expected to be about equal. The diffusion coefficient in hydrogen is nearly five times that for the above gases, which probably accounts for the more rapid initial expansion observed experimentally.

The measured values for the current densities in air are in reasonable agreement with those of Flowers (1943) and Norinder and Karsten (1949), but a close comparison is not possible, partly because of the higher currents used by these investigators, and also because the times to peak current were appreciably greater than in the present work. Flowers concludes that the channel expands until the current density falls to a practically constant value of the order of 10^3 amp/cm.; the present results tend to show that a spark carrying unit function current of between 100 and 1,000 amp. would require about 50 μsec. for the current density to fall to this value. Further expansion must occur, however, if the current is maintained, as eventually the discharge characteristics must become those of an arc, for which the current density is about 10 amp/cm^2 for currents of the order of 100 amp. (von Engel 1929).

Flowers does not differentiate in his measurements of expansion rates between the early stages and later stages of spark channel growth, and it appears that the resolution time of his camera was insufficient to record the changes occurring within the first one or two microseconds. However, Lawrence and Dunnington (1930) in their photographs of a spark carrying a current of 1,000 amp., record a cross section of about 10^{-3} cm^2 at a time 0·1 μsec. after breakdown. This corresponds to a diameter of approximately 0·36 mm. and is in reasonable agreement with the value to be expected by extrapolation from the present results, as given in Figure 7. In the studies by Raiskij (1940) a value of about 0·1 mm. is recorded for the spark diameter, but in this case the sparks were oscillatory and changes in diameter with time were not investigated. Later measurements by Raiskij (1948), concerned with the velocity of spread of the area of contact between sparks and electrodes, show that this velocity changes at the same rate as the velocity of expansion of the channel itself.

One of the interesting results of the present measurements is that the current density at a given time from breakdown is about the same for all the sparks observed, and is therefore roughly independent of the peak current, as shown by Figure 18. On the assumption that the current density remains the same in channels carrying much higher currents, the diameters of lightning channels may be estimated. For instance, with a current of 20,000 amp., which is the average magnitude of the current recorded in lightning discharges, the extrapolated value for the diameter is 2·7 cm. at a time of 10 μsec. after breakdown, and is 3·2 cm. at 50 μsec. Again, with a current of 100 kiloamp. the extrapolated diameter is 6·1 cm. at 10 μsec. and 7·1 cm. at 50 μsec. These extrapolated values are in rough agreement with the direct measurements made by Norinder and Karsten (1949), who find that the diameter of a spark channel carrying an oscillatory current, of quarter-period 12 μsec. and peak amplitude 102 kiloamp., is 8 cm. at a time of 60 μsec. after breakdown.

An estimate of the average electron concentration in the channel can be made from a knowledge of the average current density and the voltage gradient (Craggs and Meek 1946). On the assumption that the channel of cross section A cm^2 contains uniform concentrations N_+ and N_- of positive ions and electrons per cm^3, the current flowing in the channel is given by

$$I=(N_+v_+ + N_-v_-)Ae \text{ amp.,}$$

where v_+ and v_- cm/sec. are the drift speeds of the positive ions and electrons respectively, and e is the electronic charge in coulombs. N_+ and N_- must be roughly equal, or otherwise large space charge fields would be set up. As $v_+ \ll v_-$, the term N_+v_+ may be neglected in comparison with N_-v_-, so that we may write

$$I=N_-v_-eA$$

or, if i denotes the current density, in amp/cm^2,

$$i=N_-v_-e.$$

The value of the electron concentration N_- can then be computed from a knowledge of i and v_-. The value of i is given by the present experiments, while v_- depends on the voltage gradient in the channel, which is also known (Higham and Meek 1950). It may be assumed that in the short times considered (several microseconds) the gas density in the channel, and therefore the electron mean free path, is little changed from that of the surrounding gas, so that on this assumption the electron drift speed, corresponding to a particular voltage gradient in the spark channel, can be obtained by reference to electron mobility curves (Nielsen 1936). Values for N_- computed on this basis for spark channels in nitrogen at atmospheric pressure are given in Table 3. Clearly the values should be regarded as giving the order of magnitude only, in view of the approximation made by considering the current to be uniformly distributed over the channel cross section and the uncertainty regarding the applicability of electron mobility data to the different conditions obtaining in the spark channel. Again, mobility data for molecular nitrogen have been used because, although practically 100% dissociation may be expected in the channel, no data are available for atomic nitrogen. The values for N_- are of the same order as those deduced in air and other gases in previous investigations (Craggs and Meek 1946, Craggs and Hopwood 1947).

Table 3

Current wave *	500 (10)			250 (10)			250 (28)		
Time from breakdown (μsec.)	0·5	2	8	0·5	2	8	0·5	2	8
Channel diameter (cm.)	0·14	0·27	0·41	0·11	0·19	0·28	0·11	0·19	0·27
Current density (k.amp/cm^2)	32	7·9	2·2	29	7·8	2·4	29	8·9	3·7
Voltage gradient (v/cm.)	125	60	30	125	60	30	100	59	36
Electron drift velocity (km/sec.)	4	3·3	2·8	4	3·3	2·8	3·8	3·1	2·8
Electron density ($N\times 10^{-17}$/cm^3)	5	1·5	0·5	4·5	1·5	0·5	4·8	1·8	0·8

* Figures give peak in amperes, decaying to half-value in ( ) μsec.

ACKNOWLEDGMENTS

This work was commenced in the high voltage laboratory of the Metropolitan-Vickers Electrical Company Limited and the authors wish to express their thanks, particularly to Mr. F. R. Perry, for the facilities provided and for the equipment they were allowed to retain.

The prosecution of the work has been greatly stimulated by the many discussions with Dr. J. D. Craggs, to whom the authors are indebted for advice in the writing of these two papers. The authors also thank their colleague, Dr. R. F. Saxe, for making available the oscillograph used in these investigations.

One of the authors (J. B. H.) makes grateful acknowledgment to Liverpool University for a Munitions Fellowship, and to the Institution of Electrical Engineers for Scholarships which enabled the work to be continued at Liverpool University.

REFERENCES

Allibone, T. E., and Meek, J. M., 1938, *Proc. Roy. Soc.* A, **166,** 97, **169,** 246.
Beams, J. W., 1930, *Phys. Rev.*, **35,** 24.
Craggs, J. D., and Hopwood, W., 1947, *Proc. Phys. Soc.*, **59,** 755.
Craggs, J. D., and Meek, J. M., 1946, *Proc. Roy. Soc.* A, **186,** 241.
Craig, R., and Craggs, J. D., 1950, in preparation.
Dunnington, F. G., 1931, *Phys. Rev.*, **38,** 1535.
von Engel, A., 1929, *Z. tech. Phys.*, **10,** 505.
Flowers, J. W., 1943, *Phys. Rev.*, **64,** 225.
Fucks, W., and Bongartz, H., 1943, *Z. Phys.*, **120,** 468.
Higham, J. B., and Meek, J. M., 1950, *Proc. Phys. Soc.* B, **63,** 633.
Husbands, A. S., and Higham, J. B., 1950, in preparation.
Lawrence, E. O., and Dunnington, F. G., 1930, *Phys. Rev.*, **35,** 396.
Meek, J. M., 1947, *Nature, Lond.*, **160,** 110.
Nielsen, R. A., 1936, *Phys. Rev.*, **50,** 950.
Norinder, H., and Karsten, O., 1949, *Ark. Mat. Astr. Fys.*, **36A,** Pt. 4.
Ollendorff, F., 1933, *Arch. Elektrotech.*, **27,** 169.
Prime, H. A., and Saxe, R. F., 1949, *Proc. Instn. Elect. Engrs.*, **96,** Pt. II, 662.
Raiskij, S. M., 1940, *J. Tech. Phys., U.S.S.R.*, **10,** 529; 1948, *J. Exp. Theor. Phys., U.S.S.R.*, **18,** 941.
Rompe, R., and Weizel, W., 1944, *Z. Phys.*, **122,** 636.
Saxe, R. F., and Higham, J. B., 1950, *Proc. Phys. Soc.* B, **63,** 370.
Schonland, B. F. J., 1938, *Proc. Roy. Soc.* A, **164,** 132.
Suits, C. G., 1935, *Physics*, **6,** 315.
Toepler, M., 1925, *Arch. Elecktrotech.*, **14,** 305; 1926, *Ibid.*, **17,** 61; 1927, *Ibid.*, **18,** 549; 1929, *Ibid.*, **21,** 433.

The Electrical Resistance of Liquid Gallium in the Neighbourhood of its Melting Point

By C. DODD
University College, London

Communicated by E. N. da C. Andrade; *MS. received* 17*th May* 1950

ABSTRACT. In order to determine whether any discontinuous change in electrical properties takes place when a liquid is supercooled, measurements have been made on the resistance of liquid gallium. No such discontinuity has been found, the resistance varying linearly with temperature over the range investigated.

The results yield values for the resistivity and temperature coefficient of resistance of liquid gallium.

§ 1. INTRODUCTION

RECENT experiments on supercooled liquids by Dodd and Hu Pak Mi (1949) have shown for several liquids a discontinuity in the viscosity–temperature curve occurring at the melting point as the liquid passes into the supercooled region. Since liquid gallium, a liquid metal (M.Pt. 29·9° c.), readily supercools, it is of interest to see whether any discontinuity occurs in any of its properties at the melting point. Viscosity and density measurements are at present being carried out in this laboratory, and the present paper describes the results of measurements of the electrical resistance of liquid gallium in the normal and in the supercooled state.

§ 2. PROCESSING OF GALLIUM

Gallium when melted, immediately acquires a surface coating of oxide which makes the metal wet glass, a condition unsuitable for accurate resistance measurements. About 40 gm. of gallium were melted in a quartz vessel and the oxide converted into the chlorides by addition of a little dilute hydrochloric acid. The vessel was exhausted through drying tubes and the excess acid absorbed in caustic soda tubes, leaving only gallium and a surface film of its chlorides. When the quartz vessel is maintained at red heat for two hours under a reduced pressure of 10^{-5} mm. Hg, the chlorides (boiling points 210° c. and 535° c.) boil off, leaving pure gallium with a shiny surface resembling clean mercury. This gallium does not now wet glass provided the vacuum is maintained. By tilting, the gallium was poured from the quartz vessel into the bulb of the resistometer joined to it by a cone and socket. The evacuated resistometer was then sealed off near this joint using a blow-pipe.

§ 3. APPARATUS

The resistometer (Figure 1) consists of a long (45 cm.) length of precision bore (1 mm.) capillary tubing in Pyrex with 1 cm. spheres blown at each end and connected to the reservoir bulb into which the gallium is initially introduced. Side tubes are fused to the spheres, and platinum strips sealed through the glass provide electrical contact between the gallium in the spheres and the mercury in the side tubes. The whole apparatus is mounted on a frame and immersed in a well-stirred water bath with the open ends of the side tubes just above the water level. The temperature of the bath was maintained constant to better than

0·01° c. and was measured with a mercury-in-glass thermometer calibrated to an accuracy within 0·02° c.

The electrical measurements were made on a Smith Difference Bridge in which the resistance of connecting leads is eliminated, so that resistances could be measured to an accuracy of about 0·00001 ohm.

§ 4. METHOD

The resistometer was first filled with liquid gallium and its resistance determined at a series of temperatures up to 20° c. above the melting point and also in the supercooled region. Readings taken at the same temperature for different fillings did not differ by more than 0·00004 ohm.

Since the measured resistance includes not only the resistance of the gallium in the capillary but also that of the mercury in the side tubes and of the platinum contacts, an estimation of these latter two resistances was obtained by measuring the total resistance with the resistometer filled with pure mercury at 0° c. instead of gallium. The difference between this value and the value for the resistance of the mercury in the capillary tube, calculated from a knowledge of the resistivity of mercury and the dimensions of the capillary, gives the appropriate end correction at 0° c.

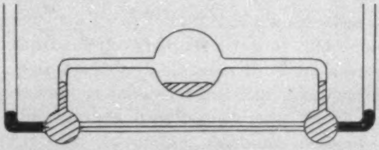


Figure 1.

The uniformity and area of cross section of the tube were estimated by the usual mercury thread method and the length of the capillary was determined by a cathetometer.

§ 5. RESULTS

The measured resistance with the resistometer filled with Ga is plotted against temperature in Figure 2 over the range 0° to 50° c. It is seen that over this range the variation is linear, and moreover there is no indication whatever of any discontinuity occurring as the liquid passes through its melting point into the super-cooled region.

This confirms the work of Bridgman (1921) who made measurements on supercooled gallium at 0° c. and found the resistance to lie on a regular prolongation of the curve for the resistance above melting point. Guntz and Broniewski (1908) on the other hand found the resistance of the liquid to pass through a minimum and to increase again in the unstable region below the melting point.

§ 6. TEMPERATURE COEFFICIENT OF RESISTANCE

Neglecting the small resistance of the gallium in the end spheres, the measured resistance with the resistometer filled with gallium at temperature t °c. is given approximately by

$$R_0 + r_0 + (R_0\alpha + r_0\beta)t$$

where R_0 is the resistance at 0° c. of the gallium actually in the capillary, r_0 is the resistance of the mercury in the side tubes also at 0° c., and α and β are the temperature coefficients of gallium and mercury respectively. The slope of the line in Figure 2 is thus equal to $(R_0\alpha + r_0\beta)$, and the intercept on the resistance axis is $(R_0 + r_0)$.

At 0° c. Measured resistance when filled with mercury $= 0{\cdot}47752$ ohm.
Calculated resistance of mercury in capillary $= 0{\cdot}47015$ ohm.

Thus $r_0 = 0{\cdot}00737$ ohm ; $R_0 = 0{\cdot}13606$ ohm ; whence $\alpha = 1{\cdot}089 \times 10^{-3}$/deg. c.

It may be noted that, like mercury, liquid gallium has a much lower temperature coefficient of resistance than most other metals for which the coefficient approximates more nearly to the theoretical value ($3{\cdot}66 \times 10^{-3}$ per degree c.).

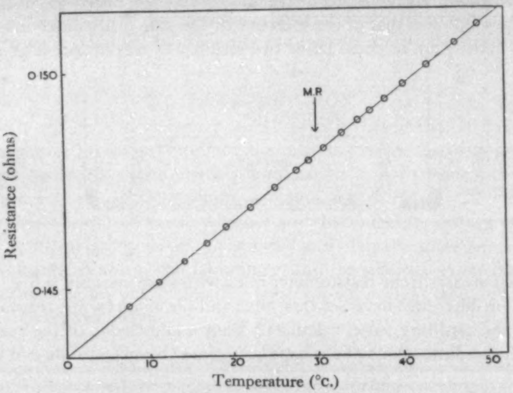


Figure 2.

§7. RESISTIVITY OF LIQUID GALLIUM AT 0° c.

The results obtained with gallium enable an estimate to be made of the resistivity of gallium at 0° c., but in view of the uncertainty of the position of the ends of the capillary tube no great accuracy can be claimed. Since however the only published values for the resistivity differ widely, it is thought that even an approximate value might prove of use.

The area of cross section and the length of the capillary tube are $9{\cdot}0438 \times 10^{-3}$ cm² and $45{\cdot}20$ cm. respectively, whence $\sigma_0 = 2{\cdot}72_3 \times 10^{-5}$ ohm. cm. This value may be compared with the value $2{\cdot}592 \times 10^{-5}$ found by Bridgman (1921) and $2{\cdot}63 \times 10^{-5}$ given by Guntz and Broniewski (1908) both these values being at the melting point.

ACKNOWLEDGMENTS

The author wishes to acknowledge his indebtedness to Dr. G. R. Davies of the Chemical Research Laboratory, Teddington, for kindly lending the supply of gallium, and to Professor E. N. da C. Andrade for the interest he continues to show in this and related work.

REFERENCES

Bridgman, P. W., 1921, *Proc. Amer. Acad. Sci.*, **56,** 109.
Dodd, C., and Hu Pak Mi, 1949, *Proc. Phys. Soc.* B, **62,** 454.
Guntz, A., and Broniewski, W., 1908, *C.R. Acad. Sci., Paris*, **147,** 1476.

Hydrogen-filled Geiger Counters

By B. COLLINGE
George Holt Physics Laboratories, The University, Liverpool

Communicated by J. R. Holt; *MS. received 27th February* 1950

ABSTRACT. Permanent gas counters are discussed and a new quenching circuit is described. Quenching times of about 45 μsec. have been used and plateaux with slopes better than 0·05% per volt obtained. A formula for the correction of counting rates for losses caused by the counter dead-time has been confirmed within wide limits. It has been found possible to measure counting rates as high as 5×10^5 counts per minute with an accuracy closer than 1%.

---

§ 1. INTRODUCTION

Since the discovery by Trost (1937) that the presence of an organic vapour considerably improves the characteristics of counters, the use of counters containing a single gas has almost ceased. This is largely because counters containing an organic vapour have self-quenching properties, and a quenching circuit external to the counter is not required. In addition, these counters have a relatively short insensitive time, enabling them to be used at counting rates higher than was possible with externally quenched counters.

The self-quenching counter has, however, a number of disadvantages. The life is limited to about 10^8 or 10^9 counts and during this time its characteristics are gradually deteriorating as the organic vapour is broken down. Also, it is found that these counters are temperature sensitive. When a high degree of stability is required, externally quenched counters filled with a permanent gas have advantages. The dead-time of such counters is electronically controlled and can have a known, constant value, thus enabling an accurate correction to be applied for losses. It is shown in this paper that such counters can be made with better dead-times and plateau slopes than can be obtained with similar self-quenching counters.

Whether or not a counter contains an organic vapour, the discharge caused by the passage of an ionizing particle is terminated by the positive ions produced during the discharge. The positive ions form a sheath around the wire which reduces the field to a value below that at which multiple ionization can occur. In the self-quenching type of counter the organic vapour prevents the production of free electrons when the positive ions are collected at the cathode. Such electrons would produce a further discharge of the counter. Counters containing a simple gas only therefore require an external quenching circuit to ensure that the voltage across the counter is less than that required for a Geiger discharge when the positive ions are collected. The reduction of the wire potential caused by the circuit must not be too great, or the time for the positive ions to move across the counter may be unnecessarily long.

During the quenching time, that is the time for which the potential is reduced by the action of the circuit, the counter is operating in the proportional region. Hence an electron produced at the cathode causes a localized avalanche at the wire. The positive ions so produced will also be collected at the cathode and there is a chance of further electrons being produced. It may be necessary, therefore, to quench for a period several times longer than the positive ion transit time, thus ensuring that no further electrons are produced in the counter after the wire potential has been restored to its full value. If γ is the probability of a positive ion producing an electron at the cathode and g is the gas amplification during the quenching time, then a condition for quenching to be effective is that γg shall be less than unity.

The time t for a positive ion to travel from anode to cathode in a Geiger counter has been shown by Stever (1942) to be given by

$$t = p\,\frac{(b^2 - a^2)}{2p_0KV}\log(b/a). \qquad \ldots\ldots(1)$$

In the case of the counter used in the present investigations a, the radius of the wire, is equal to 0·1 mm., and b, the radius of the cylinder, is equal to 0·9 cm. Substituting these values in equation (1) a value of t of about 18·3 μsec. is obtained for hydrogen at a pressure p of 28 cm.Hg, assuming a wire potential V of 2 kilovolts and a mobility K of 13·8 cm/sec/volt/cm. (Mitchell, quoted by Tyndall 1938); p_0 is the normal atmospheric pressure.

The quenching time necessary to ensure that there shall be no spurious counts must be greater than t; the actual value will depend on γ. Values of γ, at the low energies possessed by positive ions at the cathode, have not been measured directly. The work of Penning (1930), however, indicates that values of γ for diatomic gases such as hydrogen and nitrogen are small compared with the values for the ions of helium, neon and argon. Cheyney (1917) has shown that γ may be very small for low energy hydrogen ions. This conclusion is confirmed by Jones (1939) who has calculated γ from data obtained from a discharge in hydrogen using copper electrodes. Jones found that γ for hydrogen was of the order of 10^{-6} for a value of x/p equal to 20. x is the field strength in volt/cm. and p is the pressure in millimetres of mercury.

The low value of γ and the high mobility of hydrogen ions suggest that counters filled with hydrogen might operate satisfactorily with quenching times considerably less than 100 μsec. if a suitable circuit could be devised. In addition, hydrogen has other advantages as a counter gas: it is free from metastable states, it does not form negative ions by electron attachment, and counters filled with it produce larger pulses than when they are filled with other gases (Miller and Montgomery 1942).

§ 2. THE CIRCUIT

A number of circuits have been tried and found unsatisfactory at counting rates greater than a few thousand a minute. At these counting rates the counting action fails and a continuous discharge takes place. The reason for this is that the voltage on the counter returns to its full value and another discharge can occur before the circuit can be retriggered. Hence a discharge may take place which is not quenched and the chance that this may happen increases with the counting rate. An unquenched discharge will be followed by a spurious discharge which will in general have a smaller amplitude than a genuine pulse. This will be so if a discharge occurs before the field conditions in the counter are fully restored, either because the wire potential has not recovered, or because the spurious count occurred before all the positive ions had reached the cathode. If the circuit does not quench on this pulse then a second spurious discharge will occur which may be of even smaller amplitude. Unless the circuit is sensitive enough to trigger on the small pulses which constitute a continuous discharge, the counter will cease to count; thus, not only must the circuit have adequate gain but its recovery time must be less than the transit time of the positive ions.

Further, it would be desirable for a quenching circuit to have a constant quenching time independent of the counting rate to enable a simple and accurate correction for counting losses to be applied.

The circuit shown in Figure 1 fulfils the necessary conditions. It consists of a separate quenching valve V_1, the anode current of which is cut off by applying negative bias to the control grid. The counter wire is connected to the anode, which is normally at a positive potential of about 280 volts with respect to earth. The high-voltage supply for the counter is connected between the counter cylinder and earth. After amplification, by V_2 and V_4, the pulses from the counter trigger V_6 and V_8, which are connected in a 'flip-flop' circuit; the period depends mainly on the circuit components. When V_1 conducts, the anode potential falls to within a few volts of the earth potential, causing the counter wire potential to be reduced by nearly 280 volts. The quenching time is conveniently adjusted by altering the value of the resistance between the suppressor grid of V_6 and H.T. The pre-set resistances compensate for component tolerances while the potentiometer in the grid circuit of V_4 is a gain control normally set for maximum sensitivity. V_9 and V_{10} have been included to provide a negative output pulse at low impedance and may not always be necessary.

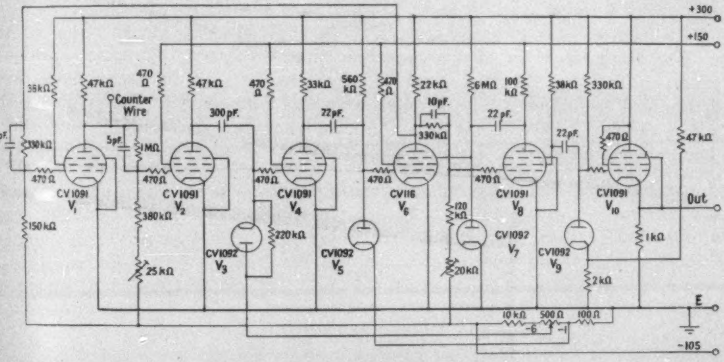

+300
+150
36kΩ
47kΩ
470 Ω
Counter Wire
47kΩ
470 Ω
33kΩ
560 kΩ
470 Ω
22kΩ
10pF.
6MΩ
100 kΩ
39kΩ
330kΩ
15pF.
330kΩ
5pF.
1MΩ
300pF.
22pF.
330kΩ
22pF.
22pF.
470Ω
47kΩ
470Ω
470Ω
470Ω
470Ω
470Ω
CV1091
V1
CV1091
V2
CV1091
V4
CV116
V6
CV1091
V8
CV1091
V10
Out
150kΩ
380kΩ
25kΩ
CV1092
V3
220kΩ
CV1092
V5
120 kΩ
20kΩ
CV1092
V7
CV1092
V9
2kΩ
1kΩ
E
10kΩ
500Ω
100Ω
-6
-1
-105

Figure 1. Counter quenching circuit.

§ 3. COUNTERS

A standard type of counter has been used in this work. It consists of a copper cathode, 12 cm. long and 1·8 cm. in diameter, mounted in a Pyrex glass tube and held in position by constrictions in the glass. A tungsten rod sealed into the glass wall and silver soldered to the cathode provides electrical contact. The tungsten wire is 0·2 mm. in diameter. Pyrex sleeves over the ends of the wire protrude about 3 mm. into the cathode to define the sensitive volume of the counter.

Each counter was washed out with concentrated nitric acid and distilled water and baked out on a vacuum system at 375° c., using a rotary pump and a diffusion pump for several hours. The copper cathode was reduced in hydrogen at 375° and then oxidized in air at the same temperature. In this way it was found possible to produce a uniform coat of oxide firmly attached to the cathode. The counter was baked out again for several hours on the pumps. After cooling it was filled with hydrogen free from oxygen and sealed off.

A pressure of 28 cm. of hydrogen was used, which gives an efficiency of roughly 99% for fast electrons (Korff 1946). Counting commenced when the potential between the counter cathode and earth was about 1,950 volts.

§ 4. MEASUREMENTS AND RESULTS

4.1. *Quenching Time*

The minimum value of the quenching time necessary to suppress spurious counts caused by electrons produced by positive ions at the cathode was determined as follows. The counter wire was connected to the anode of V_1 and a variable 2·2 kv. supply was connected between the cylinder and earth. The high-voltage supply was adjusted so that the counter was operating in the plateau region, and a series of readings of counting rate for various quenching times was obtained using a gamma-ray source to produce a suitable counting rate.

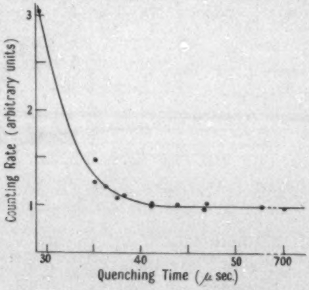


Figure 2. Counting rate as a function of the quenching time.

The results are shown in Figure 2. The counting rate remains constant as the quenching time is decreased from 600 μsec. to about 40 μsec. and it then increases rapidly at lower values.

4.2. *Counter Characteristics*

Curve (*a*) of Figure 3 is the characteristic curve of one of the standard counters; it was obtained with the circuit adjusted to have a quenching time of about 45 μsec. The flat region of the plateau is 70 volts long and has a slope of 0·03 per cent per volt. The full length of the plateau is 170 volts. The circuit abruptly ceased to function at the upper end of the plateau.

Characteristic curves were obtained with other standard counters and gave results which were similar to this curve, but the slopes of the plateaux were not quite as good.

Curve (*b*) of Figure 3 is the characteristic curve of the same counter taken after an interval of three months. The starting voltage increased by about 40 volts although the slope did not change appreciably.

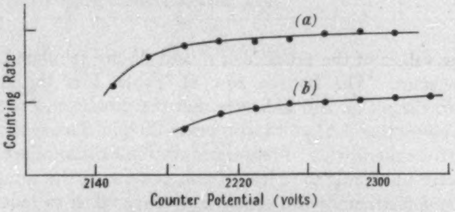


Figure 3. Characteristic curves of the standard counter. (*a*) initially ; (*b*) three months later.

A counter was made, similar to the standard counters, but with a reduced copper cathode which proved to be very photo-sensitive. It was noticed, immediately after running a continuous discharge, that the background counting rate was very high. Background counting rates of the order of a thousand per minute were observed immediately after switching the wire potential back to the operating value. The counting rate had not returned to the normal background rate after an hour. The standard counters with oxidized cathodes did not show this effect at all. It has not been found possible to suggest a satisfactory explanation.

4.3. *Circuit and Counter Tests*

The dead-time of the circuit was obtained using the paired source method. In this method two separate γ-ray sources are used and the values of the counting rates when the sources are presented to the counter, first separately and then together, are measured. The counter had the characteristic shown in curve (*a*) of Figure 3 and the readings were taken with the counter voltage adjusted to 2,280 volts. The value of the quenching time was determined from a calibrated cathode-ray oscillograph and is accurate to within about 3%.

Table 1

Quenching time (μsec.)	62·7	50·2	50·2
Dead-time (μsec.)	70·8	57·9	57·1
Difference (μsec.)	8·1	7·7	6·9

For each particular value of the quenching time ten separate determinations of the dead-time were made. The counting rate, obtained with two sources presented to the counter, varied from about 1,000 to 8,000 a second. The average value of the dead-time is shown in the second row of Table 1 for the quenching times shown in the top row. The dead-time was approximately constant within the limits of the experimental error over the range of counting rates used. The difference between these two figures is the recovery-time of the circuit and its average value is about 7·5 μsec.

The circuit and counters have also been checked using a half-gramme radium source and noting how the counting rate varied with its distance r from the counter. Each observed counting rate, N_r, was corrected for losses assuming a dead-time τ of 57 μsec., the value obtained in the previous experiment. The true counting rate N_t was obtained from the formula

$$N_t = \frac{N_r}{1 - N_r \tau}. \qquad \ldots\ldots(2)$$

In Table 2 the values of the product of r^2 and N_t are tabulated. The product should be constant. The bottom row of Table 2 is the fractional error, calculated from the mean, and indicates that the circuit and counters could be relied upon to better than 1·5 % for losses below 20 %. This is about the accuracy expected from the experiment. Measurements of the distance between the source and counter were unreliable to at least 2 cm. because of the finite size of source and counter, and furthermore no precautions were taken to reduce the effect of scattered γ-rays.

Table 2

Distance r between sources and counter (cm.)	241·6	281·6	321·6	351·6	371·6	401·6	451·6	501·6	551·6	601·6	631·6
Calculated true counting rate N_t (counts/sec.)	4645	3420	2617	2229	2002	1730	1379	1047	914	764	698
$r^2 N_t \times 10^6$	271	271	270·5	275	276	279	281	263	278	276	278
Percentage error in $r^2 N_t$	1·3	1·4	1·6	0	0·4	1·4	2·2	4	1·1	0·4	1·1

4.4. *Correction for Counting Losses*

It has been pointed out by Blackman and Michiels (1948) that equation (2) has not been satisfactorily confirmed experimentally. The apparatus described above enables an experiment to be performed to check the validity of the formula for large losses. A constant γ-ray source was used and the recorded counting rate measured for various values of the dead-time τ. The true value of the counting rate was calculated using equation (2) and the values of τ were obtained from an oscillograph which had been calibrated previously, using a pulse generator and a scale of 10^4.

The results are shown graphically in Figure 4 (*a*), (*b*) and (*c*). The calculated counting rate, obtained using equation (2), is plotted as a function of the losses expressed as a percentage of the true counting rate. The true counting rate was taken as equal to the calculated rate for a sufficiently small value of τ. Results are shown for three different positions of the source. The three curves have been combined in Figure 5. The ratio of the difference between the calculated counting rate and the true counting rate, expressed as a percentage of the true counting rate, is plotted along the vertical axis. The horizontal scale is the same as for the previous graph. The results lie on a smooth curve which should be a horizontal straight line. The extent of the deviations from this line indicates the accuracy to be expected from results obtained with the apparatus and corrected for losses using equation (2). The percentage error is less than 1 % for a loss in the counting rate of 40 %.

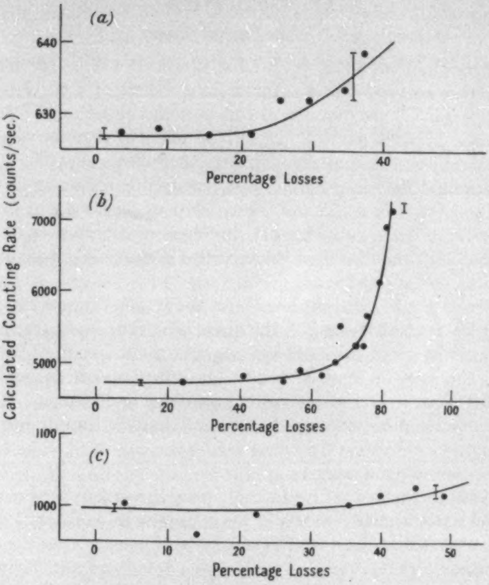


Figure 4. Calculated counting rate as a function of the percentage losses.

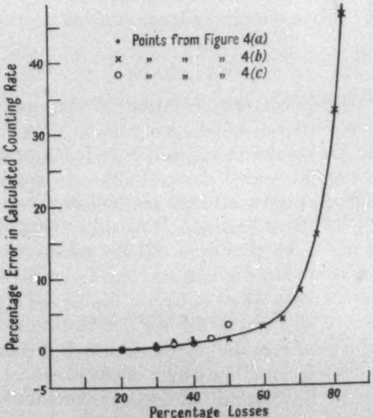


Figure 5. Percentage error in the calculated counting rate as a function of the percentage losses.

§ 5. DISCUSSION

5.1. *Ion Transit Times*

Mitchell (see Tyndall 1938, p. 70) has obtained values for the mobility of hydrogen positive ions in hydrogen for various values of x/p. His results show that for x/p less than 20 the mobility is constant and equal to 13·8 for moderately pure hydrogen. Hence equation (1) will be valid in regions where x/p is less than 20. It can be shown that x/p is greater than 20 for distances from the wire less than 1 mm. and that a positive ion will cross this region in less than 1 μsec. Thus the figure of 18·3 μsec. for the transit time of positive ions in the standard counter, calculated using equation (1), involves no serious error and should, therefore, agree with the observed value, if the disturbing effect of the positive ions on the field is neglected.

It is seen from Figure 2 that the minimum quenching time necessary to suppress spurious counts is about 41 μsec. If the quenching time is reduced below 41 μsec. then some spurious counts occur, indicating that the majority of the positive ions have reached the cathode in that time. The calculated time for one transit of the ions is 18·3 μsec., that is, two transits take place in 36·6 μsec. In the present case, therefore, quenching occurs when the quenching time is just greater than the time for two movements of the ions across the counter. The calculated and observed times are in good agreement.

Stever (1942) and others have found very poor agreement between the observed and calculated times for movements of positive ions in counters. Stever (1942) and Montgomery and Montgomery (1940) have suggested that the disagreement is due to variations of the mobility with x/p. Friedland and Krumbein (1948) have obtained better agreement by assuming a simple variation of the mobility with x/p. The results obtained here indicate that for hydrogen it is sufficient to assume a constant value for the mobility; this conclusion was reached by Hartog and Muller (1949) in the case of argon–alcohol mixtures.

5.2. *Measurement of γ*

The fact that only two transits of the ion sheath are required to ensure quenching indicates that γ is very small for hydrogen ions on oxidized copper. A rough calculation, based on the results shown in Figure 2, indicates that γ is less than 10^{-6}, in agreement with the work of Jones (1939). It may be possible to devise an experiment, similar in principle to the method used to obtain the results in Figure 2, in order to obtain values of γ. It would be necessary to measure both the number of ions in the ion sheath N_s and the gas amplification g during the time the counter is quenched.

5.3. *Length of Plateau*

It is seen from Figure 3 that the plateau is only 170 volts long, although the quenching voltage is 280 volts. The length of the plateau is determined, not by the quenching pulse, but by the potential across the counter for which the counter pulse is large enough to cause the valve V_2 to become non-conducting. The reason is as follows. Should a discharge occur which does not trigger the circuit and quench the counter, then the inevitable spurious count will not be amplified by V_2, and so quenching cannot occur. The counter then breaks down into a continuous discharge and no further counting is possible. This was confirmed by disconnecting the grid condenser of the valve V_4 and noting that V_2 became non-conducting when the counter voltage corresponded to the upper limit of the plateau.

5.4. *Calculations of True Counting Rate*

Lifschutz and Duffendack (1938) claim to have published an experimental verification of equation (2). Blackman and Michiels (1948) have pointed out that this work was not entirely satisfactory. In fact Lifschutz and Duffendack (1938) used a type I recorder which is re-excited by a further pulse occurring during the dead-time, whereas the formula only applies to type II recorders which have a constant dead-time. In the experiment described in this paper the apparatus had a constant dead-time and equation (2) is therefore applicable. From the results of Figure 5 it is concluded therefore that the formula for the correction of losses may be relied upon to better than 1% up to 40% losses.

The error caused by assuming that the dead-time of the circuit is equal to the quenching time may be neglected, because in this circuit the difference between these two times is a constant. A constant error in τ causes a constant error in the calculated counting rate, that is the curves of Figures 4 and 5 would be displaced vertically and the shape would be unchanged.

The shape of the curves can be explained by assuming that there is a small percentage error in τ which would occur if the duration of the pulse from the flip-flop circuit decreased as the on/off ratio of the circuit increases. A decrease of less than 10% in the quenching time is sufficient to explain the 50% error in the calculated counting rate at 80% losses.

No effort has been made to eliminate the small changes in the quenching time which, it is believed, cause the serious departures from the formula for losses greater than 40%. It seems unlikely that one would wish to correct for such large losses in practice.

5.5. *Performance of Hydrogen-filled Counters*

The quenching circuit used is somewhat more complicated than others described in the literature. It is believed to be superior, however, in that the quenching time can be precisely determined, thus enabling a simple and accurate correction for losses to be made up to high counting rates. It is implicit in this statement that the recovery time of the counter potential plus that of the circuit is short; the results shown in Table 1 show that it is of the order of $7{\cdot}5\,\mu$sec.

The performance of the circuit and counters can be summarized as follows: (i) adequate plateaux have been obtained using quenching times of $40\,\mu$sec.; (ii) an accuracy to better than 1% can be expected at counting rates of 5×10^5 counts per minute when the losses are about 30%, assuming a dead-time of $50\,\mu$sec.

From this it is evident that such counters would be useful when high counting rates are used over considerable periods of time and when it is necessary to have accurately reproducible conditions. The behaviour of a counter filled with a simple gas should be very reliable provided it is adequately outgassed. Such counters would be expected to have an almost unlimited counting life.

ACKNOWLEDGMENTS

The author is indebted to the Department of Scientific and Industrial Research for the grant which made the work possible and wishes to thank Dr. J. Rotblat and Dr. J. D. Craggs for many helpful discussions.

REFERENCES

Blackman, M., and Michiels, J. L., 1948, *Proc. Phys. Soc.*, **55,** 549.
Cheyney, W. L., 1917, *Phys. Rev.*, **10,** 335.
Friedland, S. S., and Krumbein, A. D., 1948, *Bull. Amer. Phys. Soc.*, p. 51, April.
den Hartog, H., and Muller, F. A., 1949, *Physica*, **15,** 789.
Jones, F. L., 1939, *Phil. Mag.*, **28,** 192.
Korff, S. A., 1946, *Electron and Nuclear Counters* (New York : Van Nostrand).
Lifschutz, H., and Duffendack, O. S., 1938, *Phys. Rev.*, **54,** 714.
Miller, C. W., and Montgomery, C. G., 1942, *Phys. Rev.*, **61,** 734.
Montgomery, C. D., and Montgomery, D. D., 1940, *Phys. Rev.*, **57,** 1030.
Penning, F. M., 1930, *Proc. K. Ned. Akad. Wet.*, **33,** 841.
Stever, H. G., 1942, *Phys. Rev.*, **61,** 38.
Trost, A., 1937, *Z. Phys.*, **105,** 329.
Tyndall, A. M., 1938, *The Mobility of Positive Ions in Gases* (Cambridge : University Press).

Cavitation produced by Ultrasonics

By B. E. NOLTINGK and E. A. NEPPIRAS
Mullard Electronic Research Laboratory, Redhill, Surrey.

MS. received 9th February 1950

ABSTRACT. The problem of cavitation produced by ultrasonic vibration is examined theoretically. Equations are developed which describe the motion of a gas-filled cavitation bubble in a liquid medium subjected to alternating pressure ; the case of an empty cavity is also considered. Information is obtained concerning the distribution of fluid pressures and velocities in the medium near the bubble surface during the motion. It is suggested that these theoretical conclusions may be used to show how the intensity of the various effects of cavitation will depend on ultrasonic frequency and intensity. In particular, it is predicted that all cavitation phenomena will diminish and finally disappear as the frequency is raised. The important part played by the nuclei from which the cavitation bubbles grow is emphasized.

§1. INTRODUCTION

In recent years, a great deal of practical work has been done concerning possible applications of ultrasonic technique to industry. Many applications are known to depend on the production of cavitation in liquids, but no thorough examination has yet been made of what exactly occurs during ultrasonically induced cavitation.

The problem of erosion caused by cavitation was first encountered in connection with rotary hydraulic apparatus, such as ships' propellers and water turbines, and it was in order to obtain some measure of the liquid pressures capable of causing erosion that the early theoretical work on cavitation was attempted. Rayleigh (1917) examined theoretically the behaviour of an incompressible fluid in which he imagined a spherical void to be suddenly formed. More recently, Beeching (1942) extended Rayleigh's analysis by taking into account surface tension effects, and the pressure of liquid vapour in the bubble. Silver (1942) introduced thermodynamic considerations, but several questionable assumptions in his treatment of the problem render his results of doubtful value.

The earlier experimental work was mainly confined to investigating the conditions under which cavitation would occur, but in 1948 Knapp and Hollander (1948) used high-speed cinematography to trace the life history of a cavitation bubble. They confirmed the existence of very large radial velocities and accelerations during the collapse period, for which the bubble motion was found to agree closely with that predicted by Rayleigh's treatment of the collapsing empty void. Quite recently, Plesset (1949) has developed an equation for the motion of a vapour-filled bubble in a changing pressure field and applied this to a complete analysis of Knapp and Hollander's experimental observations.

Usually, when cavitation is produced in hydraulic apparatus, the bubble will be at different points in the system at different stages of its life, and it is difficult, owing to the complex spatial variation of pressure, to be sure of the instantaneous pressure which is causing the bubble to alter in size. Plesset's calculations were based on the assumption that this pressure distribution was the same as that measured under non-cavitating conditions, while in Rayleigh's work, the pressure was simply taken to be constant. By contrast, in considering cavitation produced by ultrasonic waves, we need to be concerned with only one point in space, where the pressure changes are known and accurately controllable. This renders the problem amenable to more exact calculation.

§ 2. THEORY RELATING TO BUBBLE MOTION

(i) *Statement of Problem*

True cavitation (i.e. the formation of cavities or voids) in a pure liquid cannot occur until the liquid pressure has become sufficiently negative to overcome the forces of natural cohesion, and it is well known that degassed liquids can withstand very high negative pressures (Briggs, Johnson and Mason 1947, Harvey, McElroy and Whiteley 1947, Temperley 1947). But in a gas-filled liquid containing undissolved gas in the form of bubbles small enough to remain in suspension, these nuclei will expand or contract if the external pressure is made to change, and liquid will evaporate into the partial void produced. Although true voids are not present, this is the process generally referred to as cavitation when the pressure changes are sufficiently large and rapid, and it is the process we shall consider in the following theoretical discussion. As will appear more clearly below, no sharp distinction is to be expected between cavitating and non-cavitating conditions; the former is merely an extreme case.

The experimental evidence available (for example, Harvey *et al.* 1944) indicates that permanent gas nuclei are necessary for the onset of cavitation. These may, or may not, have come out of solution on solid nuclei. A third possibility exists when the liquid contains solid particles that are not wetted by the liquid; in this case, it is feasible that gas-free cavities may develop on what are essentially solid nuclei. Whatever their mode of origin, the bubbles will certainly gain vapour from the surrounding liquid as they grow, but since the pressure of vapour in equilibrium with a curved surface decreases exponentially with the curvature, and we shall generally be concerned with very small bubbles, we may neglect the pressure of vapour in the bubble in comparison with the other forces controlling the motion.

In order to examine the problem theoretically, it is necessary to make further simplifying assumptions. The liquid is taken to be incompressible and the gas content of the bubble is assumed to be constant over its life cycle. We also assume that the applied ultrasonic pressure wave is exactly sinusoidal, although the very changes in the volume of the bubble under consideration must distort the pressure wave to some extent in its neighbourhood. Again, the assumption is made that the diameter of a bubble is always much less than a wavelength.

(ii) *Equations of Motion and their Solution*

The external liquid pressure at infinity, P, can then be written $P=(P_A-P_0\sin\omega t)$ at time t, when an ultrasonic pressure wave of amplitude P_0 and frequency $\omega/2\pi$ is superposed on a pressure P_A. We wish to consider how a bubble will grow and collapse under this pressure. Let the bubble have an arbitrary radius R_0 at $t=0$ and contain gas at the equilibrium pressure (P_A+2S/R_0) where S is the surface tension of the liquid. The kinetic energy of the whole mass of liquid, density ρ, at this instant, $2\pi\rho R^3(dR/dt)^2$, can be equated to the algebraic sum of the work done by the surface tension, gas pressure and liquid pressure at infinity, giving as the energy equation :

$$\int_{R_0}^{R}\left\{4\pi R^2\left[P_0\sin\omega t-P_A+\left(P_A+\frac{2S}{R_0}\right)\frac{R_0^{\,3}}{R^3}\right]-8\pi RS\right\}dR=2\pi\rho R^3\left(\frac{dR}{dt}\right)^2$$

if the gas changes are isothermal. The differential equation of motion is found by differentiation with respect to R, giving

$$2R\left[P_0\sin\omega t-P_A+\left(P_A+\frac{2S}{R_0}\right)\frac{R_0^{\,3}}{R^3}\right]=4S+3\rho R\left(\frac{dR}{dt}\right)^2+2\rho R^2\frac{d^2R}{dt^2} \qquad \ldots\ldots(1)$$

If the nucleus consists of, not a gas sphere, but a spherical solid which is not wetted by the liquid, so that the cavity never contains any permanent gas, then $R=R_0$ when $P_0\sin\omega t=(P_A+2S/R_0)$ and the equation reduces to

$$2R[P_0\sin\omega t-P_A]=4S+3\rho R\left(\frac{dR}{dt}\right)^2+2\rho R^2\frac{d^2R}{dt^2}. \qquad \ldots\ldots(2)$$

In this case, it will be necessary for the liquid pressure to become appreciably negative before the void starts growing. These equations are insoluble mathematically, but a number of solutions were obtained on the differential analyser at the National Physical Laboratory. Examples are given in Figures 1 and 2, where $P=(P_A-P_0\sin\omega t)$ is also plotted. It will be seen that the bubble radius increases to a maximum (coordinates t_m, R_m), and thereafter shrinks with increasing rapidity until the curve is, in fact, too steep to be traced out conveniently on the differential analyser. Such integrated curves were obtained for a range of values of the parameters P_0, ω and R_0 for a water medium at a pressure of one atmosphere ($\rho=1$, $S=80$ and $P_A=10^6$ C.G.S. units). It was found by trial that wide variations in the boundary values of radial velocity dR/dt at R_0 (which is arbitrary in equation (1)) produced insignificant changes in the contour of the

(t, R) curves; it was therefore taken as zero throughout. The values used for the various parameters are listed in the Table together with the corresponding values of t_m and R_m. It will be seen later that all the information required concerning the collapse of a bubble can be readily obtained once these coordinates are known and it is interesting to see how they depend on the primary variables. The

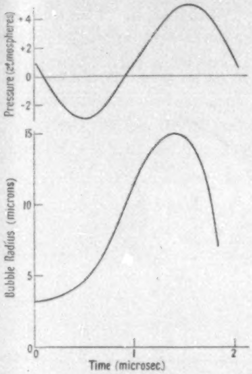


Figure 1. Radius–time curve for a gas-filled bubble.
$P_0=4\times10^6$; $\omega=3\times10^6$; $R_0=3{\cdot}2\times10^{-4}$.

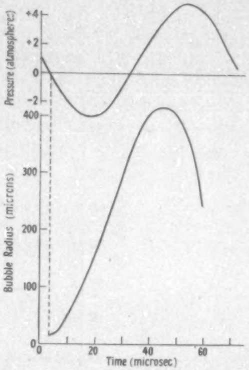


Figure 2. Radius–time curve for a void.
$P_0=4\times10^6$; $\omega=9\times10^4$; $R_0=16\times10^{-4}$.

P_0 ($\times10^6$ C.G.S. units)	ω	R_0 ($\times10^{-4}$ cm.)	R_m ($\times10^{-4}$ cm.)	t_m (μsec.)
			Obtained from solutions of equation (1)	
4	9×10^4	3·2	420	45·6
	3×10^5		128	13·7
	9×10^5		44	4·6
	$1{\cdot}5\times10^6$		28	2·8
	3×10^6		15	1·4
	9×10^6		6·7	0·5
	9×10^6	0·8	4·2	0·4
	3×10^7		1·7	0·15
	9×10^7		1·0	0·05
	9×10^4	80	496	47·3
		32	455	46·2
		16	438	46·0
		3·2	420	45·6
		1·6	412	45·3
		0·8	407	45·3
2		3·2	188	39·8
3			318	43·3
4			420	45·6
6			579	48·0
8			706	49·3
12			914	50·7
			Obtained from solutions of equation (2)	
4	9×10^6	0·8	2·4	0·4
	9×10^4	16	420	46

curves of Figures 3, 4 and 5 show them plotted against P_0, $1/\omega$ and R_0 respectively; Figure 4 is of particular interest in showing t_m and R_m to be roughly proportional to the period of the ultrasonic pressure wave over the range considered.

The very high pressures and radial velocities associated with cavitation occur during the collapse, when the walls of the bubble may be pictured as rushing inwards until they are repelled by the cushioning action of the gas in the bubble, which is by then extremely severely compressed. The total collapse time is usually a small fraction of a period of the ultrasonic vibration, so that no great errors will be introduced by regarding P as constant for its duration. The motion of the bubble is then amenable to calculation; as shown in the Appendix, the limits of bubble radius and the maximum liquid kinetic energy and surface velocity of the bubble during the collapse can easily be deduced. This latter is given by V, where

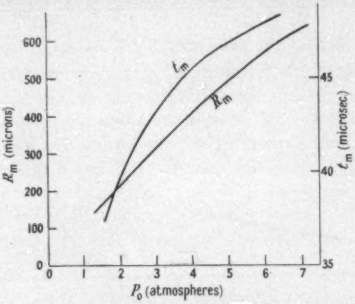


Figure 3. Variation of t_m and R_m with pressure amplitude, P_0. $\omega=9\times10^4$; $R_0=3{\cdot}2\times10^{-4}$.

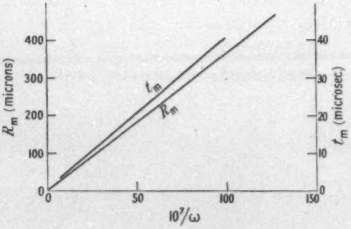


Figure 4. Variation of t_m and R_m with $1/\omega$. $P_0=4\times10^6$; $R_0=3{\cdot}2\times10^{-4}$.

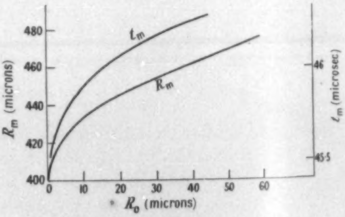


Figure 5. Variation of t_m and R_m with R_0. $\omega=9\times10^4$; $P_0=4\times10^6$.

$$V^2=\frac{2P(\gamma-1)}{3\rho\gamma}\left[\frac{P(\gamma-1)}{Q\gamma}\right]^{1/(\gamma-1)}$$

if the gas changes are taken to be adiabatic, γ being the ratio of the specific heats for the gas and Q the pressure of gas in the bubble at its maximum radius. In many of the cases considered, V is greater than the velocity of sound in water, and since it is probable that fluid velocities do, in fact, often become supersonic during the final stages of the collapse, our theory, which assumes an incompressible medium, must be thereby invalidated. However, it seems probable that deductions made from the theory, even when the liquid velocities are supersonic, can still give an indication of the relative importance of ω and P_0 in estimating the intensity of cavitation.

(iii) *Oscillating Bubbles*

We have so far only considered the behaviour of a cavitation bubble up to the instant when it collapses to its minimum size. This is obviously an unstable state, from which a series of complex oscillations would be expected to ensue. In the hydrodynamical treatment given, all forms of damping have been ignored. If these were taken into account, it is to be presumed that the amplitude of the oscillations would be found to fall off rapidly, and this is confirmed by the experimental work of Knapp and Hollander (1948).

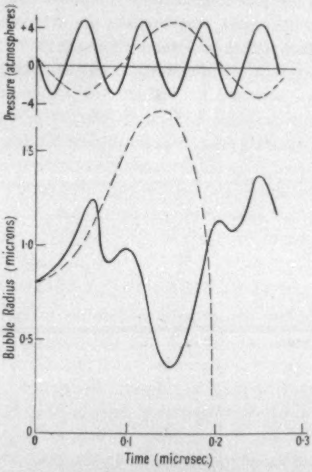


Figure 6. Radius–time curves for a gas-filled bubble at different frequencies.
$P_0=4\times10^6$; $R_0=0{\cdot}8\times10^{-4}$; $\omega=9\times10^7$ (full line), 3×10^7 (broken line).

While the events of the previous ultrasonic cycle may well be ignored if there has been time for a large number of damped bubble oscillations before the next cycle begins, matters will be different if the bubble period becomes comparable with the ultrasonic period. For $P_0 \ll P_A$, the resonant frequency of the bubble, n, is given by

$$(2\pi n)^2 = \frac{3\gamma(P_A + 2S/R_0)}{\rho R_0^{\,2}} \qquad \ldots\ldots(3)$$

(Minneart 1933), the bubble motion being approximately simple harmonic. In this case, damping of the motion is due almost entirely to sound re-radiation and heat conduction (Saneyosi 1941, Meyer and Tamm 1939, Pfriem 1940). There is no doubt that these damping terms are important at ultrasonic frequencies even when, corresponding to $P_0 \ll P_A$, the amplitude of the motion is small.

If P_0 is comparable with, or greater than, P_A, resulting in large variations in the bubble radius, the oscillations will not be even approximately simple harmonic, and in this case, even if damping is ignored, an accurate expression for the resonant radius, corresponding to (3), cannot readily be deduced. However, equation (3) may still be used to give an order of magnitude for R_0 unless $P_0 \gg P_A$. For R_0 less than the resonant figure, in which case the collapse will be completed before the end of the pressure cycle, what may be described as true cavitation will occur; for R_0 greater than the resonant value, the oscillations may be very complex, but true cavitation should not occur. This prediction can be tested, in theory, by solving the equation of motion (1) for suitable high values of ω and P_0 and a value of R_0 greater than the resonant figure calculated from (3). At 15 Mc/s. and $P_A = 1$ atmosphere, this equation shows the resonant R_0 to be about 0·4 micron. The full line curve in Figure 6 gives a solution of equation (1) taken through four complete cycles of pressure at this frequency and $R_0 = 0{\cdot}8$ micron. As can be seen, the oscillations are complex and nowhere do the radial velocities reach the high values characteristic of cavitating bubbles. The broken line curve refers to conditions identical with the above, except that the frequency has been reduced to 5 Mc/s.; as expected, it shows the steep collapse curve typical of cavitation. The inference is that for each impressed frequency there is a maximum to the size of bubbles that can give rise to cavitation. As the frequency is raised, this limit becomes smaller and smaller. It is thus apparent that, by sufficiently increasing the ultrasonic frequency, the other parameters remaining the same, the expected intensity of cavitation can be reduced to vanishing point.

§ 3. DEDUCTIONS FROM THEORY

For cavitating bubbles, the details of bubble motion and the conditions existing in the liquid medium surrounding the bubble during the collapse are dealt with in the Appendix; we now wish to show how the theoretical information obtained may be used to estimate the general intensity of cavitation action, and the way in which this varies with ω and P_0. A difficulty which immediately presents itself is the arbitrary nature of the parameter R_0; we have little knowledge concerning practical limits to the possible values of this initial bubble radius. In theory, by taking it sufficiently small, whatever values we assign to ω and P_0, we can obtain the conditions for producing cavitation pressures, etc., of almost any magnitude.* This means that the calculated magnitudes of liquid pressures, etc., obtained from the theory, may have little practical significance. However, we are primarily concerned with deciding how the intensity of cavitation will vary with ultrasonic intensity and frequency, and for that purpose it is perfectly legitimate to regard R_0 as constant, treating all our expressions as functions of ω and P_0. The phrase ‘ intensity of cavitation ’ has an obvious general meaning, but before it can be used at all quantitatively, it must be given some more exact definition. It is probable that the function which best describes the intensity will depend on the phenomenon in connection with which the cavitation is being considered.

The bubble motion during collapse is decided entirely by the three parameters R_m, P and Q, the intensity of cavitation increasing with increasing R_m and P and diminishing with increasing Q. P is governed in a simple way by ultrasonic amplitude, while Q for constant R_m depends only on R_0, being proportional to $(P_A + 2S/R_0)R_0^3$. Since R_m has been shown to be roughly inversely proportional to ω, it follows that all cavitation effects will be expected to fall off with increasing frequency, the sharpness of this fall being dependent on the power of R_m which enters into the particular expression chosen to represent the effect being considered. We can, as previously suggested, expect cavitation to disappear completely in the region of

---

* It may be noted that there is a sharp distinction here dependent on whether the nuclei are taken as gaseous or solid. For the latter, if $R_0 < 2S/(P_0 - P_A)$ no cavitation can occur.

$$\omega = \frac{1}{2\pi R_0}\left[3\gamma\left(P_A + \frac{2S}{R_0}\right)\right]^{1/2}.$$

Considering the other end of the spectrum, it appears that cavitation intensity should increase indefinitely as the frequency is reduced. It is not yet clear how far this tendency may be limited by assumptions in the theory becoming invalid. One factor will certainly tend to prevent very large values of R_m being reached —namely, the reduction in the peak negative pressure of the ultrasonic wave occasioned by the very presence of large bubbles.

The curves of Figure 7, which are drawn from the information given in the Appendix, show how the amplitude of the liquid pressure wave following a collapsing cavitation bubble rises steeply as the bubble radius diminishes, and the wave velocity increases to very large values. The parameter Z used there is R_m^3/R^3. The greatest fluid pressure during adiabatic collapse is $Q(P/3Q)^{3\gamma}$ and since $Q \propto R_m^{-3}$, this is $\propto R_m^{3(3\gamma-1)}$ or ω^{-9} approx., taking $\gamma = 4/3$ and $R_m \propto \omega^{-1}$. The graph of Figure 8 shows this maximum fluid pressure plotted against frequency for constant pressure amplitude and suitable R_0. As can be seen, these maximum pressures are enormous, but high pressures of this order of magnitude exist for only a small fraction of the total collapse time and over a very small volume of fluid near the bubble surface. In some cases the absolute value of the pressure

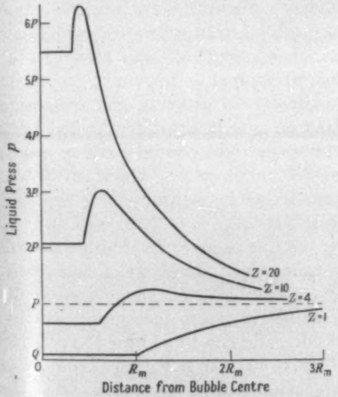


Figure 7. Pressure distribution in the fluid surrounding a collapsing, gas-filled bubble.

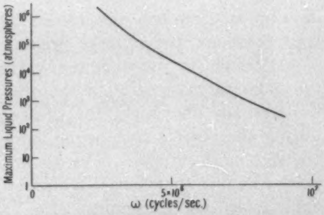


Figure 8. Variation with frequency of maximum fluid pressures during collapse. $P_0 = 4\times10^6$; $R_0 = 3{\cdot}2\times10^{-4}$; adiabatic conditions.

set up in the medium may not be so significant as the excess above some threshold figure. If this threshold can be estimated, the information given in the Appendix can be used to show over what volume, and for how long, the fluid pressures remain effective. On this basis, a comparison might be made for varying frequencies and ultrasonic intensities. Integrated functions of fluid pressures obtained in this way may, in some cases, form a useful measure of the intensity of cavitation action, but it is probable that sometimes pressure gradients will be more important than the liquid pressures themselves. As shown in the Appendix, the hydrodynamical equations give the pressure gradient dp/dr directly as a function of distance, r, at any instant of the collapse, and, in such cases, this expression will be of value in deciding the effectiveness of cavitation action.

In the special case where gas-free bubbles grow from solid nuclei in suspension in the medium, an order of magnitude for the instantaneous pressure produced on the solid particles during collapse may be obtained by a simple calculation due to Rayleigh. He assumes the nuclei to be spherical, and abandons his assumption of incompressibility for the medium at the instant that the cavity wall comes into contact with the sphere. From that instant, he further assumes that the kinetic energy of each particle of fluid is changed to elastic energy of deformation of the same particle, as determined by the volume modulus of elasticity of the fluid, E. This gives $(P')^2/2E=(\rho/2)(dR/dt)^2$ where P' is the instantaneous pressure on the surface of the rigid sphere and dR/dt the velocity of the cavity wall on impact. The instantaneous pressure on the solid surface is therefore proportional to dR/dt. On Rayleigh's assumption of a completely empty cavity of arbitrary maximum radius, these pressures will obviously be greater the smaller the nuclei, becoming infinitely great as the nuclei approach zero radius. Our theory, applied to the gasless void growing and collapsing on a rigid sphere of known radius, involves no purely arbitrary assumptions (see equation (2)), although, as already mentioned, voids cannot develop unless $R_0>2S/(P_0-P_A)$. The radius of the smallest nucleus on which voids can develop is therefore determined by P_0 and, since P_0 is limited by practical considerations, the collapse pressures are also limited.

Although, in the treatment given above, the gas changes for bubble growth are taken, for simplicity, to be isothermal and those for collapse strictly adiabatic, the true law obeyed by the gas content of the bubble will change continually during the bubble life, being more nearly adiabatic where the bubble volume changes more rapidly. The assumption of adiabatic conditions for the collapse period implies that the gas temperature will rise to very large values as the radius approaches its minimum. For a strictly adiabatic compression, the gas temperature at radius R is $T_0(R_m/R)^{3(\gamma-1)}$ and at minimum radius it is approximately $T_0(P/3Q)^{3(\gamma-1)}$, T_0 being the absolute temperature of the surrounding liquid. As an example, if $T_0=300°$ K., $P=1$ atmosphere and $Q=0{\cdot}01$ atmosphere, the gas temperature at minimum radius would be about 10,000° K. Of course, in practice, the temperature of the bubble surface will be effectively T_0, so that a steep temperature gradient must exist in the gas adjacent to it; how steep it is will depend on how nearly adiabatic the changes are. Since high temperatures are confined to the gas content, the liquid temperature remaining sensibly constant, it seems impossible that heating effects can play any important part in the destructive action of cavitation. On the other hand, the gas temperatures may rise to incandescence, so accounting for the luminescence that has been observed under ultrasonic irradiation (Harvey *et al.* 1944). Various other effects are known to be produced by intense ultrasonic waves in liquids, including emulsification, dispersion, destruction of bacteria and other organisms, and several chemical actions. It may be possible to use the cavitation formulae that have been derived in analyses of these different phenomena.

§ 4. CONCLUSION

The theoretical treatment of the cavitation problem, contained in this paper, was undertaken to show how it may be possible, on the basis of pure theory, to predict the results of ultrasonic treatment under conditions of cavitation. Equations have been derived which describe the motion of a cavitating bubble, and the information obtained from them has been used to make a few broad deductions. Exact theoretical results are not to be looked for, but certain general laws can be laid down, in particular with regard to the effect of varying the ultrasonic frequency. For comparison with experiment, further investigations are required, and a fuller development of the theory. It seems likely that high speed cinematography with the techniques at present available could give definite checks on the calculations made, and could, in particular, show the importance of damping in relation to the bubble motion. It has previously been emphasized that all results will be markedly affected by the size and number of gas nuclei present in the treated liquid. While it is difficult, on the basis of information published to date to set any useful limits to the figures to be taken for these quantities, a number of experimental approaches suggest themselves which might throw more light on the subject.

APPENDIX

Assuming adiabatic gas changes, the energy equation for the collapse of a gas-filled bubble is

$$\frac{3\rho}{2}\left(\frac{dR}{dt}\right)^2 = P(Z-1) - \frac{Q(Z-Z^\gamma)}{(1-\gamma)} \qquad \ldots\ldots(4)$$

where P, the external pressure, is taken as constant and $Z=(R_\mathrm{m}/R)^3$, R_m being the maximum radius of the bubble at the start of the collapse. In this equation, surface tension forces are ignored, since they are quite unimportant during the collapse compared with the two opposing pressure terms. For stationary values of R, $dR/dt=0$ giving $P(\gamma-1)(Z-1)=Q(Z^\gamma-Z)$ as the equation defining maximum and minimum bubble radii. This gives $Z=1$ or $R=R_\mathrm{m}$ as one limit of oscillation. At the other limit, $R \ll R_\mathrm{m}$, and $Z^{\gamma-1}=P(\gamma-1)/Q$ to a good approximation. The gas pressure corresponding to this minimum radius is

$$Q\left[\frac{P(\gamma-1)}{Q}\right]^{\gamma/(\gamma-1)}.$$

For stationary values of dR/dt, corresponding to points of inflection on the (R,t) curve, $(d/dt)(dR/dt)=0$ or $(d/dR)(dR/dt)=0$. This condition shows that, for maximum dR/dt,

$$R=R_\mathrm{m}\left[\frac{P(\gamma-1)}{Q\gamma}\right]^{1/3(1-\gamma)}$$

i.e. $R = \gamma^{1/3(\gamma-1)} R_{\text{min}} \simeq 1{\cdot}3\, R_{\text{min}}$ for air. The value of this maximum radial velocity is V where

$$V^2 = \frac{2P(\gamma-1)}{3\rho\gamma}\left[\frac{P(\gamma-1)}{Q\gamma}\right]^{1/(\gamma-1)}.$$

The total liquid kinetic energy when the bubble radius is R is

$$\frac{4\pi\rho R^3}{3}\left[P(Z-1) - Q\frac{(Z-Z^\gamma)}{(1-\gamma)}\right],$$

and this is a maximum where $(R/R_{\text{m}})^3 = Q/P$.

To investigate the distribution of fluid pressure in the medium surrounding the collapsing bubble, we apply the general hydrodynamical equation of motion. If u is the fluid velocity at distance r from the centre of the bubble, the radial acceleration of the fluid at r is

$$\alpha_r = -\frac{du}{dt} = -\frac{\partial u}{\partial t} - u\frac{\partial u}{\partial r} = \frac{1}{\rho}\frac{dp}{dr},$$

p being the pressure at r. $\partial u/\partial t$ and $\partial u/\partial r$ can be obtained as functions of r and R from the energy equation and the continuity relation $R^2\, dR/dt = r^2\, dr/dt$. We obtain

$$\frac{dp}{dr} = \frac{R}{3r^2}\left[\frac{QZ^\gamma(3\gamma-4)}{(1-\gamma)} + \frac{QZ}{(1-\gamma)} - (Z-4)P\right] + \frac{4R^4}{3r^5}\left[P(Z-1) - Q\frac{(Z-Z^\gamma)}{(1-\gamma)}\right]$$

giving, on integration,

$$p - P = -\frac{R}{3r}\left[\frac{QZ^\gamma(3\gamma-4)}{(1-\gamma)} + \frac{QZ}{(1-\gamma)} - (Z-4)P\right] - \frac{R^4}{3r^4}\left[P(Z-1) - Q\frac{(Z-Z^\gamma)}{(1-\gamma)}\right]. \qquad \ldots\ldots(5)$$

Plots of this function are shown in Figure 7 for various values of Z. To enable the curves to be drawn on a linear scale, the ratio P/Q is taken as low as 10 and Z not greater than 20. The maximum liquid pressure at any instant during the collapse occurs at a distance r_{m} from the bubble centre, obtained by solving the equation $dp/dr = 0$. The value of this maximum pressure is then obtainable from (5) by substituting $r = r_{\text{m}}$. It is easily seen that the greatest liquid pressure over the entire collapse occurs at minimum radius at the bubble surface and is equal, of course, to the gas pressure inside, QZ^γ. The pressure gradient at this point is $-QZ^\gamma/R_{\text{min}}$ to a good approximation, assuming Z to be large. Parallel equations to (4) and (5) can be derived for the case where conditions are regarded as isothermal.

The equations relating to the collapse of a completely empty cavity can, of course, be obtained from the above by putting $Q = 0$. The energy equation reduces to

$$\frac{3\rho}{2}\left(\frac{dR}{dt}\right)^2 = P(Z-1).$$

This is the case considered in detail by Rayleigh (1917). The fluid velocities, pressures and pressure gradients approach infinite values as R approaches zero.

ACKNOWLEDGMENTS

This work was carried out under the auspices of the Mullard Radio Valve Company Ltd. The authors wish to thank the directors of the company for permission to publish, and Dr. C. F. Bareford, the Manager of the Electronic Research Laboratory for his encouragement in the project. Helpful discussions with Mr. J. G. L. Michel of the National Physical Laboratory are also gladly acknowledged.

REFERENCES

BEECHING, R., 1942, *Trans. Instn. Engrs. Shipb. Scot.*, **85**, 273.
BRIGGS, H., JOHNSON, J. B., and MASON, W. P., 1947, *J. Acoust. Soc. Amer.*, **19**, 664.
HARVEY, E. N., 1939, *J. Amer. Chem. Soc.*, **61**, 2392.
HARVEY, E. N., MCELROY, W. D., and WHITELEY, A. H., 1947, *J. Appl. Phys.*, **18**, 162.
HARVEY, E. N., WHITELEY, A. H., MCELROY, W. D., PEARSE, D. C., and BARNES, D.K., 1944, *J. Cell. Comp. Physiol.*, **24**, 23.
KNAPP, R. T., and HOLLANDER, A., 1948, *Trans. Amer. Soc. Mech. Engrs.*, **70**, 419.
MEYER, E., and TAMM, K., 1939, *Akust. Z.*, **4**, 145.
MINNEART, M., 1933, *Phil. Mag.*, **16**, 235.
PFRIEM, H., 1940, *Akust. Z.*, **5**, 202.
PLESSET, M. S., 1949, *J. Appl. Mech.*, **16**, 277.
RAYLEIGH, LORD, 1917, *Phil. Mag.*, **34**, 94.
SANEYOSI, Z., 1941, *Electrotech. J., Tokyo*, **5**, 49.
SILVER, R. S., 1942, *Engineering*, **154**, 501.
TEMPERLEY, H. N. V., 1947, *Proc. Phys. Soc.*, **59**, 199.

The Colour Temperature of Light Sources

BY H. G. W. HARDING
National Physical Laboratory, Teddington, Middlesex

Lecture given to the Colour Group of the Physical Society on 4th June 1947;
revised MS. received 8th February 1950

ABSTRACT. The definition of colour temperature is explained and the approximations of the spectral distributions of energy of lamps to those of black-body radiators are indicated. The value of colour-temperature measurements to the colorimetrist is outlined and difficulties in making the measurements are dealt with. The use of photoelectric devices for colour-temperature measurements are mentioned.

§ 1. DEFINITIONS

THE International Lighting Vocabulary definition of colour temperature, with its extended meaning, states that if the standard Commission Internationale de l'Éclairage (C.I.E.) observer can colour-match the radiation from the light source with that from a full radiator (black-body radiator), then the colour temperature to be assigned to the light source is given by the temperature of the full radiator when a colour match is obtained. The light source may be a full radiator, when for a colour match there will be a spectral energy match; it may be a grey body, when for a colour match there will again be a spectral energy match, or it may be a selective radiator, when for a colour match there will probably not be an energy match.

Before going any further a full radiator, grey body and selective radiator should perhaps be defined.

A full radiator completely absorbs all radiations incident on it whatever their nature or direction. The spectral distribution of energy of the radiation emitted by it is determined solely by its temperature, and of all incandescent bodies, at the same temperature, the full radiator emits the maximum total radiation.

A grey body emits radiation the spectral distribution of energy of which bears a constant ratio at all wavelengths to that of a full radiator at the same temperature, that is, for a given temperature the colour is the same as that of the full radiator, but it is not so bright.

A selective radiator emits radiation the spectral distribution of energy of which does not bear a constant ratio at all wavelengths to that of a full radiator at the same temperature. The spectral distribution of energy of a selective radiator can, however, be almost the same as that of a full radiator at another temperature; tungsten is an example of this.

§ 2. INCANDESCENT FILAMENTS

As long ago as 1914 Paterson and Dudding in their experiments referred to the Colour Identity Temperature of platinum, tungsten and carbon lamps, and it is to light sources of this nature, that have spectral emission curves almost identical with those of grey bodies, that the term colour temperature seems to have been first applied. Evidence for believing that tungsten and carbon lamps have spectral energy distributions that are approximately the same as those of full radiators is outlined by the following experiments. Worthing (1917) showed that heated tungsten is not a grey body.

A tube of tungsten with small radial holes in it $\frac{1}{10}$ to $\frac{1}{5}$ mm. in diameter was heated and at 2,600° K. the inside and outside temperatures were within about 3° K. of each other. It had the appearance shown in Figure 1 (Plate). The broad horizontal patch represents the tube and the small bright spot the hole in it. Owing to internal reflections, black-body radiation is emitted from the hole so that it is brighter than the rest of the tube. The vertical thin line is the hot wire of the pyrometer, the brightness of which was varied to match first the tube then the spot. Measurements were made for blue light of wavelength $0{\cdot}467\,\mu$ and red light of wavelength $0{\cdot}665\,\mu$. Typical values for the ratios of the brightness of the radiation from the tungsten surface to that from the hole when the tungsten is at 2,600° K. are 0·46 for the blue light and 0·42 for the red. These values are, of course the emissivities and if the tungsten had been a true grey body the emissivities would have had the same value.

Forsythe (1923) followed on these measurements using light of three wavelengths. In effect, the temperature of the glowing tungsten was altered until the red to blue ratio was the same as that of a full radiator, then, without altering the temperatures of either the tungsten or the full radiator, the relative brightnesses for green light were measured. The ratio for the green light was found to be almost the same as it was for the red and blue. For carbon no difference was measurable. Tungsten, platinum and tantalum emitted about 1% more green light than a full radiator at 2,600° K. and osmium slightly less light. These experiments show that tungsten is not a true grey body in the sense that if it has the same temperature as a full radiator the spectral emission curve is the same shape. They show, however, that if the temperature is altered so that the colour is the same as that of a full radiator the curves, as defined for light of three wavelengths, do become almost identical in shape. Confirmatory evidence of this for light of other wavelengths is given in Figure 2.

This experimental information indicates that the lamp filaments generally used emit radiation with spectral energy distribution curves that are practically the same shape as those of a full radiator at the same colour temperature. Provided that the temperatures of the filaments are reasonably uniform and the lamp bulbs are colourless, lamps made with these filaments can be used as substitutes for full radiators.

Now, if it is known that a tungsten filament lamp has a colour temperature θ, that is, it colour-matches a black-body furnace at a temperature θ, what can be deduced from this knowledge?

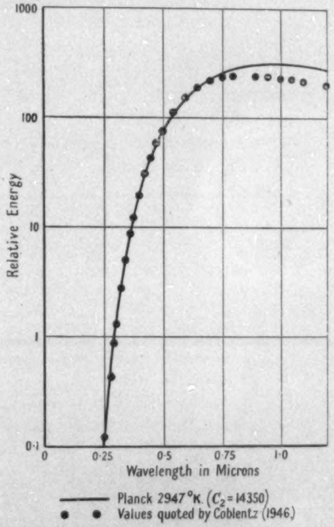


Figure 2. Spectral distribution of energy from tungsten ribbon in fused quartz bulb. *After* Coblentz, 1946. *Reproduced by kind permission of the American Astronomical Society.*

First the relative spectral energy distribution can be calculated by substituting the colour temperature for θ in Planck's formula

$$E_{\lambda\theta} = \frac{C_1\lambda^{-5}}{\exp\left(C_2/\lambda\theta\right)-1}. \qquad \ldots\ldots(1)$$

The only unknown for the evaluation of the relative distribution is C_2, which for colorimetric purposes has generally been taken as being equal to 14,350 micron degrees. The calculation is laborious as six-figure accuracy is generally required for the preparation of standard tables. Makowski (1949) of the Admiralty Research Laboratory, Teddington, has made a slide rule which gives between two and three figure accuracy. This rule is exceptionally useful for experimental work as it gives a great deal of information in addition to the relative energies.

Having the relative energy distribution, calculated by substituting the colour temperature for θ in Planck's formula, the chromaticity of a lamp can be calculated from this energy distribution and the standard C.I.E. distribution coefficients for the equal energy spectrum (Harding and Sisson 1947), and this chromaticity can then be plotted on the chromaticity chart. This procedure is adopted in colorimetry for tungsten filaments lamps and these lamps provide a wide range of light sources of known chromaticities and spectral emissions.

§3. C_2 IN PLANCK'S FORMULA

With reference to the value of C_2 in Planck's formula it should be mentioned that in 1948 the Comité International des Poids et Mesures (C.I.P.M.) adopted a modification of the international high temperature scale that involved a change in the accepted value of C_2. This change affects the colour-temperature scale and through this the properties of standard illuminants for colorimetry defined by the C.I.E. in 1931. Methods of dealing with the problems created are at present under consideration by the appropriate bodies here and elsewhere.

The fact that the relative spectral distributions of energy depend only on the ratio C_2/θ helps to a limited extent in estimating the effects of changes in the value assigned to C_2. If an energy distribution has been calculated for a temperature θ with a value C_2, then if C_2 is altered to C_2' this same distribution will serve for another temperature θ' given by

$$C_2/\theta = C_2'/\theta'. \qquad \text{......(2)}$$

§4. ESTIMATION AND MEASUREMENT OF COLOUR TEMPERATURES

4.1. *Direct Method*

The measurement of colour temperature of the lower temperature range is straightforward and is illustrated by Figure 3 (Buckley, Collier and Brookes 1924).

A full-radiator furnace, a photometer and a comparison lamp are set up as shown. The lamp is colour-matched with the furnace and then a standard lamp is placed on the furnace side of the photometer and is matched against the comparison lamp. The temperature of the furnace is measured with the pyrometer which views it through the prism. The basis for the furnace temperature scale (International Temperature Scale C.I.P.M. 1948) is the melting point of gold, Planck's radiation formula with C_2 equal to 14,380 micron degrees and the absolute temperature as −273·15° c. The gold melting point, determined by gas thermometry, has the internationally fixed value of 1,063° c. or 1,336·15° K. The pyrometer measurements are brightness–temperature measurements and are made with a red filter over the eyepiece so that the pyrometer only measures the brightness of the furnace compared with the brightness of gold at its melting point, for red light. Furnace measurements were made up to 2,400° K. for the present N.P.L. scale, but higher temperatures have been used both here and abroad. Tungsten filament lamps will operate at temperatures up to about 3,640° K., when the filaments melt. The extension of the scale up to this temperature using real furnaces is not possible at present and it is not desirable that standard lamps should be calibrated at the higher temperatures because their life will be so short. Indirect methods have therefore to be used in order that the standard lamps shall operate at comparatively low temperatures when a reasonable life can be expected from them.

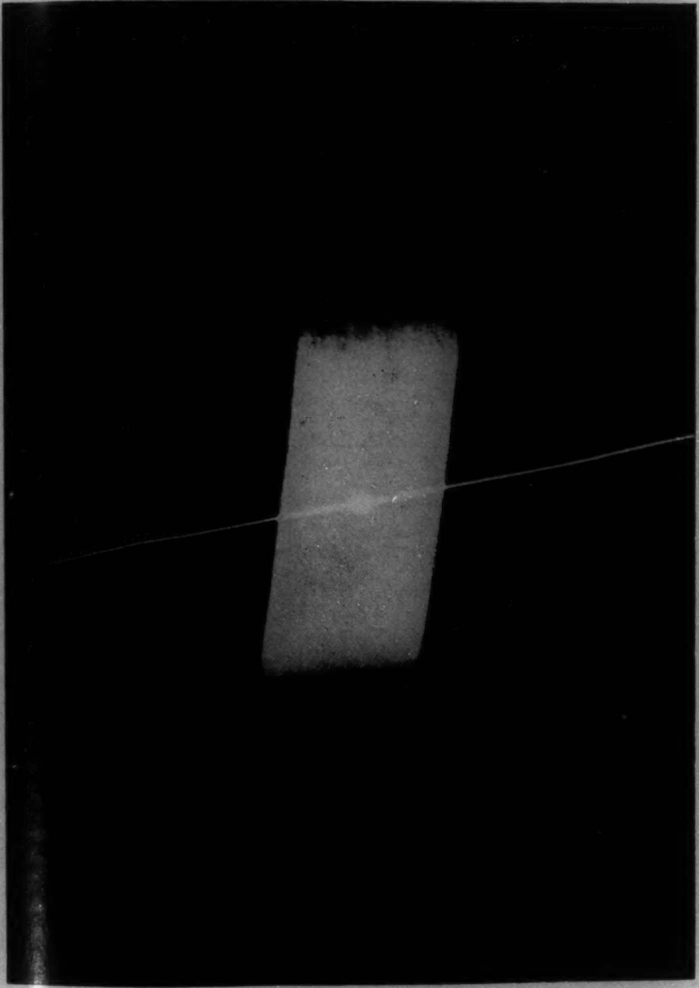


Figure 1. Photograph showing tungsten tube black-body with pyrometer filament projected against small radial hole.
Reproduced by kind permission of the International Commission on Illumination, New York.

To face page 688

4.2. *Filter Method*

The method of extending the scale beyond 2,400° K. at the N.P.L. is to use the lamps calibrated against a furnace up to 2,400° K. together with a coloured glass filter. The filter can be yellow or blue in colour. If it is yellow (Harding 1944) it is placed in front of the lamp at the higher colour temperature, if it is blue it is placed in front of the lamp at the lower colour temperature. The filter has to be carefully made and its calibration is exacting because, for instance, an error of 1% in the spectral transmission for light of one wavelength may cause an error of 1° K. in the calculation of its colour temperature step.

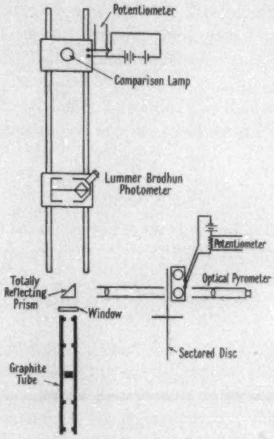


Figure 3.

Reproduced by kind permission of the International Commission on Illumination, New York.

The fact that a single filter can be used for the whole temperature scale can be illustrated by considering two lamps set up on a photometer bench with an ideal filter in front of one of them. The spectral distribution of the energy from the lamp with the filter in front of it is given by

$$E_{\lambda\theta_1} = \frac{t_\lambda C_1 \lambda^{-5}}{\exp(C_2/\lambda\theta_1) - 1}, \qquad \ldots\ldots(3)$$

where t_λ is the spectral transmission of the filter, C_1 and C_2 are constants, λ is the wavelength of the light from the radiator and θ_1 is the lamp colour temperature.

The spectral distribution of energy from the bare lamp is

$$E_{\lambda\theta_2} = \frac{C_1 \lambda^{-5}}{\exp(C_2/\lambda\theta_2) - 1}. \qquad \ldots\ldots(4)$$

Now for a colour match, energy match and intensity match at the photometer these spectral energies must be identical so that

$$t_\lambda \frac{C_1 \lambda^{-5}}{\exp(C_2/\lambda\theta_1) - 1} = \frac{C_1 \lambda^{-5}}{\exp(C_2/\lambda\theta_2) - 1}. \qquad \ldots\ldots(5)$$

The shape of the spectral transmission curve is therefore given by

$$t_\lambda = \frac{\exp(C_2/\lambda\theta_1)-1}{\exp(C_2/\lambda\theta_2)-1} \simeq \exp\left\{\frac{C_2}{\lambda}\left(\frac{1}{\theta_1}-\frac{1}{\theta_2}\right)\right\}. \qquad \ldots\ldots(6)$$

The omission of the -1 in Planck's formula means that Wien's formula is considered to be a good enough approximation to Planck's formula in this instance because for tungsten lamp temperatures $\exp(C_2/\lambda\theta)$ will always be greater than about 200. For actual filters the highest value for t_λ is arranged to be about 90%.

One filter then made with a spectral transmission curve to this specification will convert the spectral emission of a lamp at one temperature to that of a lamp at another temperature and it will do this whatever the temperature of one of the lamps, because temperature only occurs in the form $(1/\theta_1 - 1/\theta_2)$. The value of this, which is constant for each particular ideal filter, is called the reciprocal filter step. To save calculating reciprocals too frequently Priest (1933) introduced a new term, the Mired, which is a million divided by the colour temperature in degrees Kelvin, so that if a filter has a step of 100 mireds we know that

$$\frac{10^6}{\theta_1\,^\circ\text{K.}} - \frac{10^6}{\theta_2\,^\circ\text{K.}} = 100. \qquad \ldots\ldots(7)$$

There is one other fact to be got from the formula for the spectral transmission. Taking logarithms for equation (6)

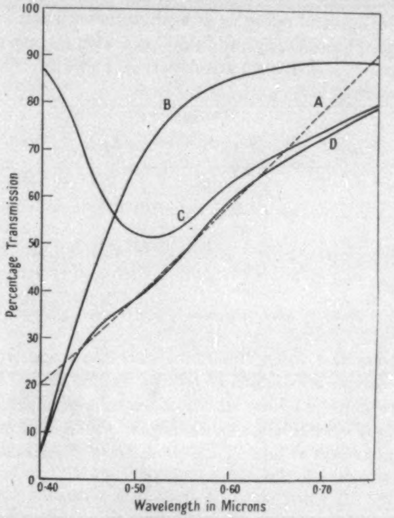
$$\log\frac{1}{t_\lambda} = \frac{C}{\lambda}\left(\frac{1}{\theta_2}-\frac{1}{\theta_1}\right)\log e. \qquad \ldots\ldots(8)$$

As $\log(1/t_\lambda)$ is equal to the optical density of the filter the reciprocal filter step is proportional to the density, so that, for instance, two similar filters together would have double the reciprocal filter step of one alone. The reciprocal step of an ideal filter is approximately proportional to its thickness; this is not strictly true because reflection from the glass surfaces should be considered as well as the approximation of Wien's to Planck's formula.

The calculation of the spectral transmission of the filter is straightforward, but the realization of such a filter is difficult because it involves accurate spectrophotometry, accurate glass working and the reasonable selection of the coloured components. For work when one or two degrees accuracy is aimed at, at least two coloured components are required. One component generally has a spectral transmission curve about the required shape for the energy match and the other is generally a pale purple or green constituent to improve the colour match. Liquid filters are often of little value because of their lack of stability, diffusion at the container surfaces after a short time, and high temperature coefficients. Glass filters are generally made and even these have to be reconditioned by repolishing the surfaces after about five years because self-blooming alters the spectral transmission curves by a few tenths of one per cent and this can be sufficient to alter the colour by an amount that will introduce observer differences of several degrees. The properties of yellow filters (Harding 1944) made at the laboratory are illustrated by the next two figures.

Figure 4 shows the approximation of the actual to the ideal spectral transmission curve and Table 1 the colour-matching properties. With this filter the chromaticities of the filtered light are less than a fiftieth of the smallest perceptible difference from the Planckian chromaticities shown in brackets.

The equipment required to make the filter that has been considered is not generally available and it is unlikely that many potential users of such filters could get them. During the war there was a demand for filters of this kind and the Admiralty asked Messrs. Chance Brothers to produce large-scale melts of a blue filter that would do for approximate colour temperature and intensity measurements. The N.P.L. indicated the required colorimetric properties and examined the various glass melts that were made. The result was Chance OB9 glass (Harding 1948). The original melt of glass was about twice as good for its purpose as any



A Theoretical Filter. B Corning 346 2·490 mm. C Corning 507 0·370 mm; D Actual Filter.

Figure 4.

Table 1

Temperature of filter (° C.)	Illuminant (° K.)	Colour quality	Percentage transmission	Equivalent colour temperature (° K.)
20	3200	0·47664 X+0·41366 Y+0·10970 Z (0·47663 X+0·41364 Y+0·10973 Z)	52·2	2497·5
20	2848	0·49708 X+0·41526 Y+0·08766 Z (0·49708 X+0·41525 Y+0·08767 Z)	52·8	2279·7
20	2550	0·51680 X+0·41456 Y+0·06864 Z (0·51679 X+0·41454 Y+0·06867 Z)	53·4	2086·5
20	2360	0·53064 X+0·41269 Y+0·05667 Z (0·53063 X+0·41265 Y+0·05672 Z)	53·9	1958·8
31	2848	0·49729 X+0·41527 Y+0·08744 Z (0·49728 X+0·41526 Y+0·08746 Z)	52·6	2277·6

other glass that we have examined and a subsequent melt has been even better. OB9 glass is not as good as the two-component filter made at the laboratory, but this could scarcely be expected as the improvements were discontinued as soon as the original colorimetric specifications for a filter for approximate measurements were satisfied.

The ratios of the spectral energies at a photometer when a lamp at 2,850° K. (E_{2850}) is matched for chromaticity and luminance with a lamp at 2,250° K. having an OB9 filter in front of it (OB9$_{2250}$) are shown in Table 2.

Table 2

Wavelength	OB9$_{2250}$/E_{2850}	Wavelength	OB9$_{2250}$/E_{2850}	Wavelength	OB9$_{2250}$/E_{2850}
0·40	0·89	0·54	0·97	0·68	1·16
0·42	0·97	0·56	1·04	0·70	1·31
0·44	1·01	0·58	1·04	0·72	1·32
0·46	1·01	0·60	1·01	0·74	1·30
0·48	0·98	0·62	0·97	0·76	1·29
0·50	0·97	0·64	0·90		
0·52	0·98	0·66	0·96		

Reproduced by kind permission of the Institute of Physics.

The residual greenness when colour matches are made with OB9 glass and tungsten filament lamps are indicated by the values of Δy in Table 3. The Δy value has been suggested by the author as a means of expressing the greenness of filters as compared with Planckian radiators. If x_1 and y_1 are chromaticity coordinates for a filter and x_1 and y_2 are those for a Planckian radiator that has the same coordinate x_1 as the filter, then $\Delta y = y_1 - y_2$.

Table 3. Values of Δy for given Thicknesses of OB9 Glass

Colour temperature of illuminant (° K.)	Δy				
	Thickness of OB9				
	1 mm.	2 mm.	3 mm.	4 mm.	5 mm.
1500	+0·0008	+0·0018	+0·0028	+0·0036	+0·0038
2000	+0·0010	+0·0019	+0·0021	+0·0015	+0·0000
2500	+0·0008	+0·0012	+0·0006	−0·0006	−0·0021
3000	+0·0006	+0·0005	−0·0003	−0·0014	−0·0028
3500	+0·0004	+0·0001	−0·0007	−0·0018	

Reproduced by kind permission of the Institute of Physics.

(The reciprocal filter stop of OB9 glass is about 60 mireds/mm.)

A Δy value of 0·001 indicates a just perceptible chromaticity difference, as observed in a photometer with a ten degree Lummer-Brodhun field. It can be seen from this table that for the thicknesses of glass that are likely to be used for most colour-temperature measurements there will never be more than a just perceptible chromaticity difference.

Priest's method of measuring colour temperature might be mentioned here (Priest 1922) because it is essentially a filter method. A quartz plate of suitable thickness was placed between Nicol prisms, as illustrated in Figure 5. This arrangement produced a filter of a variable spectral transmission which is characteristic of the ideal filter that has been considered previously, namely $t_{\lambda} = Ae^{B/\lambda}$, where A is an arbitrary constant and B is proportional to the difference of the reciprocal temperatures and to C_2. Priest's equipment did not give this precisely. By the use of this equipment Priest estimated that the colour temperature of tungsten at its melting point was 3,644° K. and that of a carbon arc with solid carbons was 3,780° K.

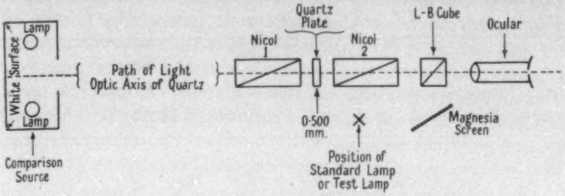


Figure 5.

After Priest, 1922. *Reproduced by kind permission of the Optical Society of America.*

4.3. *Selective Radiators*

There are two reasons for considering the problem of selective radiators. The first is that it is often convenient to assign a colour temperature to light sources such as fluorescent lamps and artificial daylight fittings so that a single number is used to describe the chromaticity of the light. These light sources do not emit radiation with spectral energy distribution curves having the same shapes as those of full-radiator furnaces, and their chromaticities are very rarely identical with those of black bodies, so that some means of assigning colour temperatures has to be sought. A measurement of their chromaticities is a much more satisfactory procedure. The second reason is that when filters are made for colour-temperature measurements the chromaticity of the lamp and filter in combination is not always identical with that of a Planckian radiator. The colour-temperature steps of the filters have to be estimated for light sources at various colour temperatures and this cannot be done unless colour temperatures can be assigned to filter chromaticities which do not quite coincide with those of Planckian radiators.

Consider what is done when an attempt is made to use colour temperature to grade the chromaticities of illuminants. The chromaticities of the full radiators are plotted as well as the chromaticity of the light source for the standard C.I.E. observer. If the chromaticity happens to be on the full-radiator locus, the temperature of the full radiator that has this chromaticity can be read from the plot and so the equivalent colour temperature is obtained. The fact has, however, to be faced that if the estimate was put to a practical test and if the spectral emission of the light source was vastly different from that of the full radiator many real observers might not agree with the estimate to within the least measurable chromaticity difference.

If the point representing the chromaticity coordinates does not come on the full-radiator locus the position is worse. A point on the locus has to be chosen that will identify the chromaticity. This is done by applying the results of experiment to a calculation and the success of the assignment of a colour temperature is limited by observer differences.

Effects of small chromaticity differences.

The results of some of our experiments illustrate the effects of small chromaticity differences on colour temperature determinations. Our first experiment was to see what would happen if an attempt was made to colour-match a tungsten filament lamp with another of almost the same energy distribution, but with a slightly greener colour. The apparatus consisted of two amps on a photometer bench and an almost colourless glass filter resembling a piece of window glass. The lamps were first colour-matched at 2,400° K. then the piece of glass was placed in front of one of them and the voltage across the other was altered in an attempt to obtain a colour match between the unfiltered lamp and the filtered one. The results of this experiment are illustrated by Figure 6.

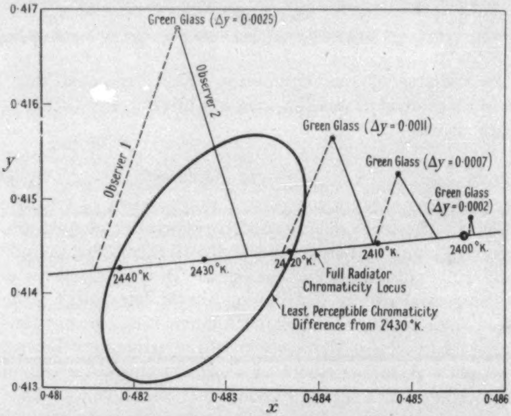


Figure 6. Observers' estimates of colour temperature for chromaticities not coincident with those of full radiators.

Observer 1 altered the colour temperature of the unfiltered lamp much more than observer 2 when the chromaticity difference was about the maximum tolerable. As the chromaticity differences were decreased by using different glass filters, the observers decided to estimate the best colour match in a certain way, which is indicated by drawing a line of defined gradient through the point representing the chromaticity to meet the full radiator chromaticity locus. These lines are often called iso-temperature lines (Judd 1936, MacAdam 1943). It is interesting to see that the observers maintain their method of estimation even when no chromaticity difference is perceptible. This experiment was repeated on different days and the observers repeated their previous settings to within about 2° K. and still maintained their 20° K. difference. The difference appears to be independent of the type of glass used to produce the greenness, that is the greenness may be produced by reducing the red or the blue light.

This observer difference has been put to the practical use of rejecting lamps that are slightly green. The two extreme observers of a group determine their observer differences for a filter of known greenness; they then compare lamps with bulbs known by previous measurement to be almost colourless with those to be inspected and if there is no observer difference the lamps are satisfactory, if there is one, the observer difference gives a very approximate measure of the greenness.

The next experiment was to see what happens when there is a calculated colour match for the C.I.E. observer, but when the energy match is slightly different. Where the energy match is within about 5% over most of the visible spectrum and within about 20% at the red and blue ends observers were found to agree to better than 5° K.

The third experiment was to show how observers reacted to small chromaticity differences when an effort was made to measure the colour-temperature steps of a slightly imperfect filter by means of lamps calibrated for colour temperature. The results that are quoted refer only to an OB9 blue glass filter with a colour-temperature step of about 2,250 to 2,850° K., but the procedure outlined can be used for any unknown filters. Two lamps and a Lummer-Brodhun photometer with a ten-degree field were set up on a bench. One lamp was operated at 2,250° K. and the blue filter was placed in front of it. The voltage across the terminals of the other lamp was varied until the best colour match was obtained. It was found that whilst each observer could repeat his settings to about 2° K. the observers differed by about 15° K. These results are illustrated in Figure 7.

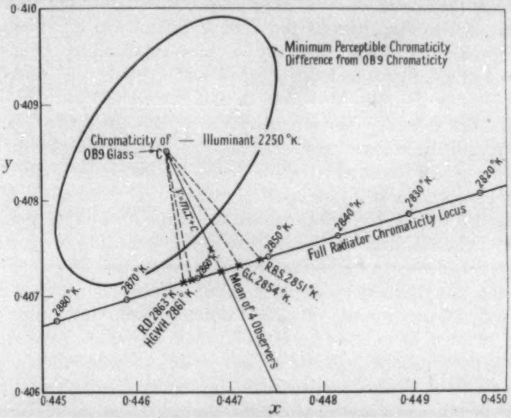


Figure 7. Gradient of the line to be drawn to estimate the colour temperature corresponding to the chromaticity of Chance OB9 glass.

4.4. *Selective Absorption Methods*

The methods of measuring colour temperature that have so far been considered have compared the whole spectral range of visible radiation from a light source with that from a standard lamp, and the measurements have therefore been made at the colour of some full radiator. If a light source had been slightly green this would have been noticed and steps taken to deal with the greenness, for instance the chromaticity may have been measured with a colorimeter because it was considered that a colour temperature could not be assigned to it. There are methods in which the whole of the spectral range of the visible light emitted by the light source is not considered which may be quite satisfactory for routine determinations, but care needs to be taken with them. They generally depend on measuring the ratio of the intensity of the red light to that of the blue light from a source by letting the light fall on either a photoelectric cell or the eye, first when a red filter is placed in front of the source, then when a blue filter is placed there. Calibrated lamps are generally chosen as standards for determining the red to blue intensities for known colour temperatures. This method is satisfactory when the energy distributions are Planckian. If the energy distributions are not Planckian the results may have very little reference to the colour temperature, for instance, the method would give a red to blue ratio for a mercury or a sodium or a daylight lamp but the chromaticities corresponding to full radiators at these colour temperatures would be appreciably different from the actual chromaticities of the light sources. For light sources with a Planckian distribution this method can be made an absolute one if the spectral transmissions of the filters and the spectral sensitivity of the photoelectric cell are known, because the ratio of the red to blue responses can then be calculated for Planckian radiators at various temperatures. Photoelectric cell sensitivities are, however, difficult to measure with an accuracy that would ensure that the final result would be more accurate than a visual measurement made by a direct method in which the whole visible spectral range of light was used. Some errors can be avoided if the red to blue ratios are measured for one temperature within the range and the calculated scale then corrected according to the result, but a safer method is to use a comparison method with a standard lamp. The precision of these methods is generally about 10° K. in colour temperature for a one per cent change in red to blue response ratio, and different lamp bulbs may give results differing by as much as 10° K. at 2,800° K. from the visual comparison.

The design of some photoelectric devices for red to blue ratio methods might be mentioned here. An attempt should be made to eliminate the effects of fluctuations of the intensity of the light source. This generally means that two cells have to be used, one covered with the red and the other covered with the blue filter. The cells are connected in opposition and a meter arranged to detect out-of-balance currents. The colour-temperature scale is provided by the positions of shutters arranged to vary the amount of light falling on the cells to obtain equality of output. If high precision is required, and especially if rectifier cells are used, it is desirable to arrange for the intensity of the incident radiation to be kept at a predetermined level. It is also desirable that all the incident light should pass through a filter having a high transmission for visible radiation, but falling to about 10% at $0{\cdot}40\,\mu$ and $0{\cdot}70\,\mu$. This restricts the response of the cells to the visible spectrum range which is desirable for colour-temperature measurements and it also reduces the drifts that sometimes occur when red light falls on rectifier cells. The red and blue filters are chosen for their high transmissions for red or blue light. If the transmission band can be restricted to the red and blue regions of the spectrum only, the maximum sensitivity of the arrangement will be obtained. This cannot generally be achieved with actual filters, but fortunately the sensitivity is not greatly diminished if the filters do transmit some light in other spectral regions. If very dense red and blue filters are used and if the colour temperature of the illuminant is altered by a specified amount, the out-of-balance cell current will only be small because very little light will reach the cells. This arrangement would give the maximum red to blue response ratio, because the spectral bands transmitted could be those near to the red and blue ends of the spectrum. Alternatively, if the filters are barely coloured, both clear glass for instance, each cell current will be large, because of the intensity of the incident light will be hardly reduced, but the out-of-balance current would be zero because the change of colour temperature of the illuminant will have the same effect on each cell if the spectral sensitivities are similar. There is therefore some value for the filter transmissions that will give maximum sensitivity. Calculations at the laboratory have shown, for example, that for filters of the types described used with rectifier cells the transmissions as measured with the cell and isolating filter should be between 25 and 50%.

A comparator made at the laboratory using selenium rectifier photoelectric cells with a galvanometer of sensitivity 150 mm/μa. has a precision of about 2° K. for the tungsten filament lamp range. In this instrument the photoelectric cells are mounted side by side behind a hole in a sphere and shutters are arranged to cover the cells by varying amounts. One cell is covered with Chance OB3 and the other with Chance OY1 glass. OB3 glass is apparently no longer available, but Chance OB8 or OB2 glasses should provide a satisfactory alternative. Light enters the sphere through a hole in the wall opposite the cells and a disc at the centre of the sphere prevents direct light from falling on them. This entrance hole is covered with a filter made from Chance OY9 and Corning 978 glasses for isolating the visible spectrum, and an annular cell mounted just in front of the hole and behind the filter is used as an indicator for the intensity of the incident light. Chance OY9 is no longer available, but OY18 should provide a reasonable substitute.

A small portable two-cell colour-temperature meter with the same glass filters and a 15–0–15 μa. meter has a range from 1,800° K. to infinite temperature marked on a scale 2 cm. long. The sensitivity is about 100° K. at 2,000° K. and about 1,000° K. at 10,000° K. In order to calibrate this photoelectric meter for daylight colour temperatures a visual instrument has been made. This uses an OB9 glass wedge with auxiliary filters and the Δy values of the chromaticities of the filters are always less than 0·0025 for the colour-temperature range from 2,000° K. to infinite temperature.

Red to blue ratio methods are particularly useful when the colour temperature of a light source, like a tungsten filament lamp, has to be determined when the lamp is in a sphere coated with a paint that is a good diffuser but is not strictly neutral in colour. If the light source is interchanged with a calibrated lamp the colour temperature of this standard lamp can be varied until the red to blue ratio is the same as that for the unknown light source and so the colour temperature of the unknown light source can be determined. The neutrality of the paint, the spectral sensitivity of the photoelectric cell and the spectral transmissions of the filters need not be known: all that has to be determined is the red to blue ratio for both light sources when they are in the sphere.

In another type of instrument which has been used for colour-temperature measurements some of the light from the source passes through a yellow filter and some through a dichroic filter which transmits red and green light. The ratio of the red to green intensities from the dichroic filter is altered by a colour wedge until the light from the yellow filter colour-matches that from the dichroic filter. Results obtained with this type of instrument depend on the observer's colour vision and an initial individual calibration is necessary if large errors are to be avoided.

4.5. *Indirect Methods*

One method, requiring expensive equipment, is to let the light from a lamp pass through a monochromator and be detected either visually or photoelectrically. If the lamp has been standardized for colour temperature when it is operated at a particular voltage the colour temperature at any other voltage can be obtained by changing the voltage to that required and determining the ratio of the energies emitted for radiation of two or three wavelengths, then finding by calculation from Planck's formula the temperature of the radiator that would give these ratios. Lamps can also be compared in this way, but it is advisable that the sensitivity of the equipment should be great enough for measuring the light after it has been allowed to fall on a diffusing medium such as magnesium oxide.

Another method of extending the colour temperature scale of a lamp, if its calibration is known for a few points, is to plot the logarithm of the power in watts against the logarithm of the colour temperature. The result is a straight line that can be extended beyond the known calibration. An experiment showed that if the colour temperature of a 500-watt or 1,000-watt lamp was determined for 2,000, 2,200 and 2,400° K. the 3,000° K. point could be estimated with an error of about 5° K. The values of the gradients of the lines joining the log (watts) and log (colour temperature) points for the 500-watt and 1,000-watt lamps were within 1% of each other.

§ 5. CONCLUSIONS

Colour temperature is very useful to express the operating conditions of light sources such as tungsten filaments, but should not be used indiscriminately for other light sources, because, if the colour and energy producing the colour are different from those of a full radiator, observer differences and minimum perceptible colour differences have to be considered before the value of colour temperature to specify the colour can be assessed.

ACKNOWLEDGMENT

Some of the work described above has been carried out as part of the general research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory.

REFERENCES

Buckley, H., Collier, L. J., and Brookes, F. J. C., 1924, *Sixth Session of the Commission Internationale de L'Éclairage* (Cambridge : University Press, 1926), p. 203.
C.I.P.M., 1948, *Échelle Internationale de Température de* 1948, *Procès Verbaux des Séances du Comité International des Poids et Mesures*, t. XXI, 1948.
Coblentz, W. W., 1946, *Amer. Astr. Soc.*, **10**, 222.
Forsythe, W. E., 1923, *J. Opt. Soc. Amer.*, **7**, 1115.
Harding, H. G. W., 1944, *Proc. Phys. Soc.*, **56**, 21; 1948, *J. Sci. Instrum.*, **25**, 333.
Harding, H. G. W., and Sisson, R. B., 1947, *Proc. Phys. Soc.*, **59**, 814.
Judd, D. B., 1936, *J. Opt. Soc. Amer.*, **26**, 421.
MacAdam, D. L., 1943, *J. Opt. Soc. Amer.*, **33**, 18.
Makowski, M. W., 1949, *Rev. Sci. Instrum.*, **20**, 876.
Paterson, C. C., and Dudding, B. P., 1914–15, *Proc. Phys. Soc.*, **27**, 230.
Priest, I. G., 1922, *J. Opt. Soc. Amer.*, **6**, 27; 1933, *Ibid.*, **23**, 41.
Worthing, A. G., 1917, *Phys. Rev.*, **10**, 377.

Certain Properties of Electrostatic Fields Encountered in Electron Lenses*

By P. A. LINDSAY
Imperial College, London†

MS. received 8th February 1950

ABSTRACT. The equipotential lines in a bipotential electron lens exhibit a fine structure which is revealed by application of the relaxation method to the solution of the Laplace equation in cylindrical coordinates. The asymmetry of the potential field between the two edges of the cylinders is confirmed by the same method and is established to be of the order of 2% in one particular case.

§ 1. INTRODUCTION

It appears from some publications on electron optics (Cosslett 1946, Zworykin *et al.* 1945, Motz and Klanfer 1946) which have dealt with the potential distribution in a bipotential lens, that there are two details which, while probably familiar to many workers in electron optics, have not been stated explicitly. This is possibly due to considerable mathematical difficulties connected with the solution of electrostatic field equations for actual boundary conditions encountered in electron optics.

§ 2. FINE STRUCTURE OF THE EQUIPOTENTIAL LINES

During the use of the relaxation method (Southwell 1946, Allen *et al.* 1945) for theoretical investigations of the potential distribution in a bipotential lens consisting of two coaxial cylinders of equal diameter D and situated $D/8$ apart it was noticed that the equipotential curves change the sign of their curvature twice in the region between the axis of symmetry and the walls of the cylinders. Since no previous theoretical results indicated the existence of such apparent irregularity the results arrived at by the application of the relaxation method might have been questioned originally on the grounds of inaccuracies connected with the transition from the differential to the finite difference equations. However, this apparently unusual shape of the equipotential curves is clearly noticeable in the experimental results published by Maloff and Epstein (1938, p. 74). They do not mention this explicitly, possibly only due to the fact that the order of quantities involved just corresponds to the order of errors common in the method of electrolytic trough measurements. The electric potential function has been calculated in the present case to an accuracy within approximately 0·1%, which exceeds by almost two orders the accuracy of most previously published results (e.g. Bertram 1942). It was further established that increasing the number of points at which the value of the potential function was calculated and hence the overall accuracy of the results, indicated more and more clearly the existence of this double change in the sign of curvature of the equipotentials in the region under consideration. Figure 1 shows the equipotential lines as calculated by the relaxation method (full lines) compared with those obtained by applying Bertram's (1942) linear approximation.

---

* This paper forms part of a thesis submitted to the University of London for the degree of Ph.D.
† Now with Electric and Musical Industries Ltd., Hayes, Middx.

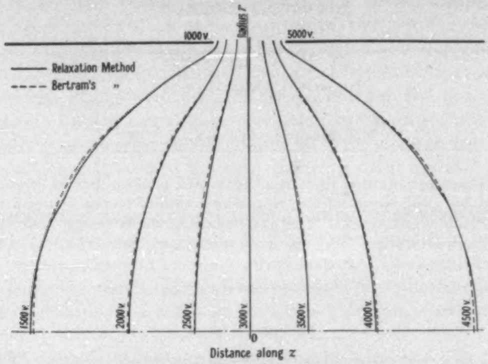


Figure 1. Equipotential lines.

It might be of interest to add that further investigations of the shape of the equipotentials in a bipotential lens indicate that the double change in the sign of curvature disappears in the limiting case of an infinitely small gap between the cylinders.

§ 3. GENERAL PROPERTIES OF THE POTENTIAL FUNCTION BETWEEN THE EDGES OF THE CYLINDERS

The other point of interest concerns the general properties of the potential function in a bipotential lens. It is seldom sufficiently emphasized that, taking the arithmetic mean of the cylinder potentials as a reference point, the potential function along the line connecting the inner edges of the two cylinders is not in general skew symmetrical. This can be proved easily by considering the additive properties of the potential function and constructing simple boundary conditions, which would add up to give actual potentials (e.g. 1,000 v. and 5,000 v.) applied across the cylinders.

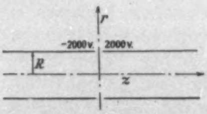


Figure 2. Boundary conditions in case I.

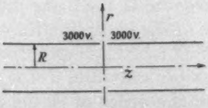


Figure 3. Boundary conditions in case II.

First consider two cylinders of equal but opposite potentials of $-2{,}000$ v. and 2,000 v. as in Figure 2. Then for reasons of symmetry the equipotential line corresponding to the arithmetic mean of the cylinder potentials (in our case 0 v.) must coincide with the r axis, and the potential function between the two edges along the surface $r = R$, R being the cylinder radius, must be skew symmetrical relative to the plane $z = 0$, as shown in Figure 4.

Next apply across both cylinders a potential of 3,000 v. as shown in Figure 3. This time the potential function along the surface $r = R$ must be symmetrical relative to the plane $z = 0$, because the conditions are identical on both sides of the plane $z = 0$. Yet the value of the potential function anywhere along $z = 0$ is less than 3,000 v. (at least as long as we assume that $\phi \to 0$ for $r, z \to \infty$). Adding cases I and II of Figure 4 in order to obtain 1,000 v. and 5,000 v. boundary conditions, it is obvious that the potential function between the edges of the two cylinders consists in general of the sum of a symmetrical and a skew symmetrical function and hence its value would differ at $z = 0$ from the arithmetic mean of the cylinder potentials. Consequently the equipotential corresponding to this value of the potential will not be the plane of geometrical symmetry $z = 0$, but some other surface, slightly concave towards one of the cylinders, as shown in Figures 1 and 4.

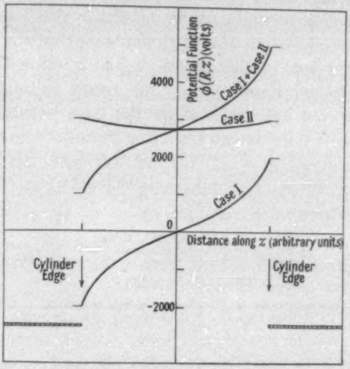


Figure 4. Potential distribution for two different boundary conditions.

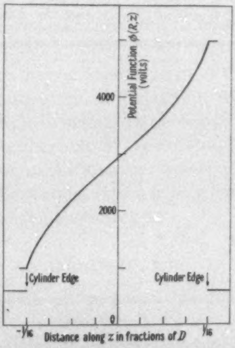


Figure 5. Potential distribution in the gap between cylinders.

z in fractions of D	$-\frac{1}{16}$	$-\frac{3}{64}$	$-\frac{1}{32}$	$-\frac{1}{64}$	0	$\frac{1}{64}$	$\frac{1}{32}$	$\frac{3}{64}$	$\frac{1}{16}$
$\phi(R, z)$ in volts	1000	1733	2199	2579	2932	3291	3588	4190	5000

Using ordinary analytical methods, it seems impossible to calculate the shape of the potential function in the space between the edges, because all the more drastic approximations usually introduced in the calculation of the potential distribution in a bipotential lens have to be applied of necessity to this particular region. Yet using the relaxation method and allowing for smaller accuracy around the edges of the cylinders it is still possible to show the general properties of the potential function in that part of the lens. The results support previous expectations and indicate clearly the non-symmetrical properties of the function. Making no claims to great accuracy of the results in this mathematically difficult region of the lens, it was shown in one particular case that the function at the point half way between the edges of the cylinders had a value that was short of the arithmetic mean by 2% (see Figure 5). This result indicates clearly that if we assume the function to be symmetrical in the gap between the edges of the cylinders, as is frequently done in extrapolating boundary conditions, as suggested by Cosslett (1946), Zworykin *et al.* (1945) and Bertram (1942), uncertainty of the resulting potential distribution calculations is at least of this order of magnitude.

REFERENCES

ALLEN, D. N. DE G., SOUTHWELL, R. V., and VAISEY, G., 1945, *Proc. Roy. Soc.* A, **183**, 258.

BERTRAM, S., 1942, *J. Appl. Phys.*, **13**, 496.

COSSLETT, V. E., 1946, *Introduction to Electron Optics* (Oxford : Clarendon Press).

MALOFF, I. G., and EPSTEIN, D. W., 1938, *Electron Optics in Television* (New York : McGraw-Hill).

MOTZ, H., and KLANFER, L., 1946, *Proc. Phys. Soc.*, **58**, 30.

SOUTHWELL, R. V., 1946, *Relaxation Methods in Theoretical Physics* (Oxford : Clarendon Press).

ZWORYKIN, V. K., *et al.*, 1945, *Electron Optics and the Electron Microscope* (New York : John Wiley and Sons).

Reduction of the Spherical Aberration of Magnetic Electron Lenses

By U. F. GIANOLA
Department of Electron Physics, University of Birmingham

Communicated by J. Sayers; *MS. received* 18*th January* 1950

ABSTRACT. The dependence of the resolving power of the asymmetrical bell-shaped magnetic field on the lens parameters is examined for varying degrees of field asymmetry. It is shown that a high asymmetry index favours high resolution, and a preferable value of lens power is specified.

By suitably reinforcing the lens field with that of a small air-cored coil, the theoretical resolving power of the lens is increased. On account of the small size of the coil, the coil current must be pulsed intermittently if an appreciable reinforcing field is to be produced. The basic factors governing the design and operation of a suitable coil are discussed.

---

§ 1. INTRODUCTION

The spherical aberration inherent in high resolution electron lenses sets an ultimate limit to the resolving power of the electron microscope. Extensive researches, both theoretical and experimental (Scherzer 1947, Gabor 1949, and others), have not yet yielded practical means of correcting this aberration.

Nevertheless, it is theoretically possible to reduce the spherical error below any desired limit by a suitable choice of the form of the imaging field (Rebsch 1938). Unfortunately, the magnetic saturation of the pole-pieces sets a practical limitation to such improvements in the magnetic lens (Cosslett 1946, Liebmann 1946, Glaser 1949), although Cosslett has shown that an appreciable gain can be obtained by an appropriate increase in accelerating voltage and lens dimensions.

The possible diminution of the aperture defect by the introduction of localized fields, produced by subsidiary coils, into existing lenses will be considered in the following for certain idealized instances.

§ 2. THE LIMIT OF RESOLUTION OF MAGNETIC LENSES

The axial fields typical of practical magnetic lenses may be closely approximated by the asymmetrical bell-shaped fields which have been the subject of a rigorous analytical treatment by Glaser (1940) and Dosse (1940), who gave, in particular, the spherical and chromatic aberration constants, C_s and C_c respectively, in terms of the lens power parameters k_a and k_b :

$$k_a^2 = ea^2H_0^2/8mU; \quad k_b^2 = eb^2H_0^2/8mU. \qquad \ldots\ldots(1)$$

Here $2a$ and $2b$ are the widths of the two component symmetrical distributions at half-field strength, H_0 is the central field strength, and U is the accelerating voltage.

Dosse (1940) has evaluated these aberration constants for the object position giving high magnification, for field asymmetries of degree $q = a/b = k_a/k_b$.

In the case of the Glaser distribution, the minimum error (Glaser 1949) may be expressed in terms of the independent lens parameters :

$$d_0 \simeq [C_s(U, b, H_0)\lambda^3(U)]^{1/4} \quad \text{for} \quad \alpha = \alpha_0 \simeq \lambda/d_0 \; ; \qquad \ldots\ldots(2)$$

λ is the de Broglie wavelength of the illuminating electrons.

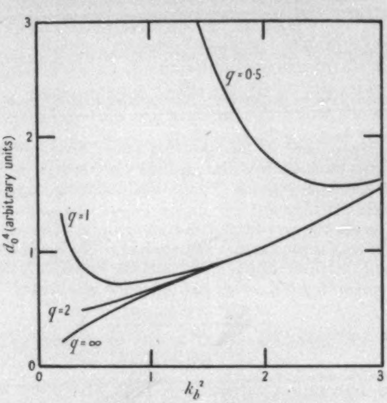


Figure 1. Variation of the minimum error with lens power at constant field strength and half-field width.

H_0 and b constant; $k_b^2 \propto 1/U$; $d_0^4 \propto C_s k_b^3$.

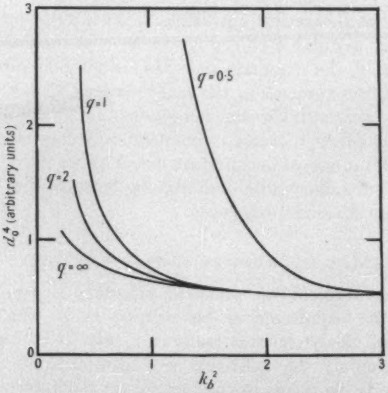


Figure 2. Variation of the minimum error with lens power at constant accelerating voltage and field strength.

U and H_0 constant; $k_b^2 \propto b^2$: $d_0^4 \propto \dfrac{C_s}{b} k_b$.

The specific dependence of this error function on each of these parameters individually, under high magnification conditions, is illustrated in Figures 1, 2 and 3.

It will be noted that, quite generally, the minimum error is smaller for field distributions of large asymmetry factor q, although the gain over that of the symmetrical field (i.e. $q=1$) is small.

For distributions of small asymmetry factor, that is of the order of unity and less, an optimum value of voltage exists for predetermined values of b and H_0 (Figure 1), as has been shown previously (Cosslett 1946, Liebmann 1946). However, for large asymmetry factors no such optimum appears for practical values of lens power, the error decreasing progressively with accelerating voltage.

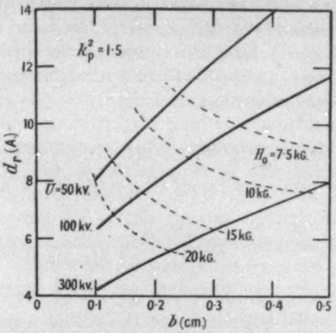


Figure 3. Variation of the ultimate limit of the resolving power of the bell-shaped field electron lens with half-field width at specific accelerating voltages. The power of the lens is kept at the preferred value $k_p{}^2 = 1\cdot5$; the loci of specific field strength values are also shown. The chromatic error is calculated here for a voltage ripple of $2\cdot5\times10^{-5}$, and an initial energy of thermionic emission of 1 ev.

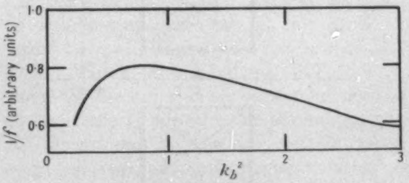


Figure 4. Variation of refractive power with the field distribution half-width only. U and H_0 constant.

A further increase in resolution is obtained when the half-field width only is increased (Figure 2). Since it is desirable to keep the dimensions of the electron-optical system within practical limits, it is preferable to adjust the lens power to keep the focal length of the objective small. As is illustrated in Figure 4, a minimum focal length is obtained at an optimum value of the half-field width. However, the fractional variation over a range of lens power is so small, for example $(f)_{k^2=2} = 1\cdot17(f)_{k^2=0\cdot8}$, so that this factor may be neglected. Any increase in lens power above about $k^2 = 1\cdot5$ does not result in an appreciable diminution of the minimum error. Hence, in the following the value $k_p{}^2 = 1\cdot5$ will be adopted as a preferential lens power. In addition, for this preferential value of lens power, the small differences in the characteristics of different asymmetric distributions may be neglected.

A more rapid decrease in the error function with change in lens power results when the field strength only is varied, as is illustrated in Figure 3, which shows the decrease in the total error d_r (Zworykin *et al.* 1945) consequent on a decrease in the half-field width, accompanied by a corresponding increase in the field strength to maintain the lens power at the preferred value k_p^2. The loci of the appropriate field strength parameter are given for specific values.

It follows from these considerations that the resolving power of existing lenses may be further improved at a determined voltage, by an appropriate increase in field strength and a reduction in lens dimensions. Unfortunately, as already indicated (Cosslett 1947), this course is restricted by magnetic saturation of the pole-pieces. Nevertheless, some progress in this direction may still be made by the introduction of subsidiary coils into existing lenses, or by the use of an air-cored coil alone as an electron lens.

§ 3. THE MAGNETIC FIELD PRODUCED BY AN AIR-CORED COIL

The magnetic field produced by an air-cored coil of axial symmetry may be expressed :

$$H(0, z) = 2\pi i \int_{-\infty}^{\infty} dz_0 \int_{r_1(z_0)}^{r_2(z_0)} \frac{r^2}{[r^2 + (z - z_0)^2]^{3/2}}\, dr, \qquad \ldots\ldots(3)$$

where i is the current through unit cross section of the coil, assumed uniform, and r_1 and r_2 are the radial boundaries of the coil at a particular axial point z_0.

For simplicity only coils of rectangular section (Figure 5) will be considered here. The axial field strength is then given by :

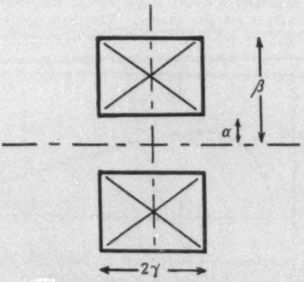


Figure 5. Cross-sectional diagram showing the dimensions of the air-cored coil considered.

$$\left.\begin{aligned} H(0, z) &= 2\pi i\alpha[k_1\eta(h, k_1) - k_2\eta(h, k_2)] \\ \eta(h, k) &= \ln\left[\frac{h + (h^2 + k^2)^{1/2}}{1 + (1 + k^2)^{1/2}}\right] \end{aligned}\right\} \qquad \ldots\ldots(4)$$

where $h = \beta/\alpha$; $k_1 = (z + \gamma)/\alpha$; $k_2 = (z - \gamma)/\alpha$.

This field may be treated in accordance with the Glaser distribution for a suitable choice of coil dimensions.

In the following it will be assumed that the axial fields H_1 and H_2, produced by the main lens and a subsidiary coil, may be represented by the symmetrical

Glaser distributions of half-widths $2a_1$ and $2a_2$, and that the respective radial axes are coincident. The resultant axial field may then be written

$$h_z = \frac{h_0}{1+s}\left[\frac{1+s+u(r+s)}{(1+u)(1+ru)}\right], \qquad \ldots\ldots(5)$$

where $r=(a_1/a_2)^2$; $s=H_2(0,0)/H_1 0,0)$; $u=(z/a_1)^2$.

It is assumed here that the influence of the subsidiary field on the magnetic saturation of the main field is small. This is justified if the relative dimensions of the coil are small.

An examination of equation (5) shows that the resultant axial field distribution will be given by that of equation (1) only if $a_1/a_2=1$, in which case $(h_z)_{r=1}=(1+s)H_1$. However, the resultant field may in general be considered in terms of the Glaser distribution to a reasonable approximation. It follows that, by reinforcing the main lens with a suitable subsidiary field, or by using an air-cored coil alone, a resultant field of increased magnitude and small half-width may be effected, and, in accordance with the above considerations, an improvement in resolution at a given voltage should be obtained.

In the absence of analyses of the electron-optical properties of other field distributions, the results of other possible combinations of main and subsidiary fields will need to be dealt with by numerical integration of the ray equation, or by direct experiment.

§ 4. PRACTICAL CONSIDERATIONS

The practical dimensions of air-cored coils needed to produce the compact fields characteristic of high power electron lenses are necessarily small. The average coil radius should be of the order of the required field half-width b, with cross-sectional area of the order of b^2. Thus the construction of a suitable coil without introducing lens asymmetries would not prove easy. Fortunately, such asymmetries may be compensated by the use of correcting cylindrical elements (Scherzer 1947, Hillier and Ramberg 1947).

In addition, in view of the small cross section of the coil, it would not be possible to use steady currents to produce the several thousand gauss necessary to reduce the lens aberrations by an appreciable factor. However, by intermittently pulsing the coil current, magnetic fields of any reasonable magnitude may be obtained, and may indeed exceed the main lens field strength by a considerable factor. In practice, fields of the required form up to 100,000 gauss have been obtained without appreciable difficulty.

Assuming that adiabatic conditions prevail during the pulse, for an allowable temperature rise ΔT in the coil the maximum pulse length τ may be written

$$\tau = \frac{J\rho\sigma}{\mu}\frac{\Delta T}{i^2}, \qquad \ldots\ldots(6)$$

where ρ, σ, and μ are the density, specific heat and resistivity of the coil winding, J is the electrical equivalent of heat, and i is the current density in E.M.U./cm².

An allowable temperature rise of 150° C. in an enamel-coated copper wire coil gives the approximate relation

$$\tau \sim \frac{2}{i^2} \times 10^6 \text{ second,}$$

with a maximum repetition frequency ν

$$\nu \sim \frac{3S}{(\beta^2-\alpha^2)\gamma^2} \times 10^{-3},$$

where S denotes the total coil surface area active in dissipating heat.

As a practical example, a coil of dimensions, $2\alpha = 0{\cdot}2$ cm., $2\beta = 0{\cdot}32$ cm., $2\gamma = 0{\cdot}12$ cm. will produce a field with a half-width b of 1 mm. This coil may be made of 180 turns of 46 S.W.G. enamelled copper wire; then a 20 amp. current pulse will produce an axial field maximum of 20,000 gauss, and the maximum pulse length is 500 μsec. with a repetition frequency of 5–10 pulses per second.*

Pulsed operation of electron lenses would severely increase the complexity of a microscope system, and, more important, would result in a reduced average image intensity and in correspondingly longer exposure times, with the existing techniques of image observation and recording. Consequently, the probability of image displacement, due to vibration and thermal drift, would be increased. However, an appreciable improvement in the efficiency of electron-image conversion should follow on using long afterglow luminescent screens coupled with an image amplifier. Pulsed operation of the electron source will also result in an intensity increase.

Other difficulties arise from the transient thermal expansion of the coil and from transient astigmatisms. Thus, in view of the many complications, the introduction of pulsed lenses may not prove feasible. However, considerable complications also accompany other methods suggested for the reduction of the aberrations of electron lenses (Scherzer 1947, Zworykin *et al.* 1945). Thus, in spite of the inherent practical difficulties, an air-cored lens on the lines indicated may serve a useful function in the development of electron microscopes of higher resolving power.

ACKNOWLEDGMENTS

The author is indebted to Professor J. Sayers for his encouragement throughout this work, which was made possible by a grant from the Department of Scientific and Industrial Research.

REFERENCES

COSSLETT, V. E., 1946, *Proc, Phys. Soc.*, **58,** 443; 1947, *J. Sci. Instrum.*, **24,** 40.

DOSSE, J., 1940, *Z. Phys.*, **117,** 316.

GABOR, D., 1949, *The Electron Microscope*, 2nd edition (London: Electronic Engineering).

GLASER, W., 1940, *Z. Phys.*, **117,** 285; 1949, *Acta Phys. Austriaca*, **3,** 38.

HILLIER, J., and RAMBERG, E. G., 1947, *J. Appl. Phys.*, **18,** 48.

LIEBMANN, G., 1946, *Phil. Mag.*, **37,** 677.

SCHERZER, O., 1947, *Optik*, **2,** 114.

REBSCH, R., 1938, *Ann. Phys., Lpz.*, **31,** 551.

ZWORYKIN, V. K., *et al.*, 1945, *Electron Optics and the Electron Microscope* (New York: John Wiley).

---

* Investigation has shown that stabilization of the current pulse to 1 in 10^6 should not be difficult.

The Computation of Wave-Front Aberrations of Oblique Pencils in a Symmetrical Optical System

By W. WEINSTEIN
Technical Optics Section, Imperial College, London

MS. received 19*th October* 1949 *and in amended form* 23*rd January* 1950

ABSTRACT. Aberration formulae given previously, together with new ones for astigmatism and secondary spherical aberration, are put into forms which are suitable for computation and simplify the treatment of field lenses. A method of comparing the formulae with ray tracing is given, and the results for a triplet objective are discussed.

---

§ 1. INTRODUCTION

In a previous paper (Weinstein 1949, referred to here as I) the author gave equations for the aberration coefficients of a pencil round a principal ray at a finite field angle, the coefficients being those for aberrations which vary as the fourth and third powers of the aperture. To these can be added equations for the secondary axial spherical aberration coefficient (see § 2 below), this being an approximation to the sixth power spherical aberration coefficient for oblique pencils. The total of results obtained, namely sagittal and meridian focal distances, third and fourth power oblique aberration coefficients and sixth power axial aberration coefficient, gives an approximate indication of the aberrations of an oblique pencil.

In order to test this, the aberration coefficients of a number of optical designs were calculated and the wave-front aberrations predicted by the coefficients were compared with those calculated from ray-tracing results, these being obtained by numerical integration of the transverse aberrations of various skew and meridian rays of the pencil.

For purposes of computation convenient formulae were derived from the fundamental aberration equations given in I and in § 2 below; also a method of computing astigmatism along a principal ray was used which involves tracing only an s-fan accurately, the quantity $1/t-1/s$, which is of aberrational magnitude, being then computed to, say, three significant figures.

These computation formulae and some of the numerical results are given in this paper.

§ 2. AXIAL SECONDARY SPHERICAL ABERRATION

In this section refraction and transfer formulae for axial secondary spherical aberration are given.

Take rectangular axes with the origin O at the pole of a refracting surface, Ox along the axis of the system, and Oz perpendicular to the axis. Then the equation of the section by the (x, z) plane of a wave front having primary and secondary spherical aberration may be written

$$x=\frac{1}{2l}z^2+\left(\frac{1}{8l^3}+C_4\right)z^4+\left(\frac{1}{16l^5}+C_6\right)z^6+O(z^8), \qquad \ldots\ldots(1)$$

where l is the radius of curvature of the wave front at the point O and C_4 and C_6 are coefficients of primary and secondary spherical aberration. When C_4 and C_6 vanish equation (1) reduces to that of a sphere of radius l, neglecting quantities which are of the order of z^8, so that the optical path difference between this sphere and the wave front is $N\{C_4z^4+C_6z^6+\mathrm{O}(z^8)\}$, where N is the refractive index of the medium; this expression is therefore the spherical aberration at semi-aperture z, referred to the paraxial focus.

The refraction and transfer formulae for C_4 and C_6 can be derived by the methods used in I.

Refraction and transfer formulae for C_4:

$$\Delta(NC_4)=\tfrac{1}{8}\Delta\left\{\frac{N}{l}\left(\frac{1}{r}-\frac{1}{l}\right)^2\right\}, \qquad \ldots\ldots(2)$$

$$C_4^{(+1)}=\left(\frac{l'}{l^{(+1)}}\right)^4 C_4'. \qquad \ldots\ldots(3)$$

Refraction and transfer formulae for C_6:

$$\Delta(NC_6)=\Delta\left\{N\left\{\frac{5}{2}\left(\frac{1}{r}-\frac{1}{l}\right)^2-\frac{3}{r}\left(\frac{1}{r}-\frac{1}{l}\right)+\frac{1}{r^2}\right\}C_4+\frac{N}{16l^2}\left(\frac{1}{r}-\frac{1}{l}\right)^3\right\}, \qquad \ldots\ldots(4)$$

$$C_6^{(+1)}=\left(\frac{l'}{l^{(+1)}}\right)^6\left\{C_6'+\frac{8l'd'}{l'-d'}C_4'^2\right\}. \qquad \ldots\ldots(5)$$

In these formulae Δ denotes the change in a quantity on refraction, primes denote quantities after refraction and the superscript $(+1)$ denotes the quantity just before refraction at the next surface ; r is the radius of curvature of the refracting surface and N and N' the refractive indices of the two media.

The derivations of equations (2) to (5) will not be given here since (2) and (3) are equivalent to the ordinary primary aberration formulae, while a formula equivalent to (4) and (5), but using different variables, has been given by von Rohr (1920, p. 240).

§ 3. ASTIGMATISM ALONG A PRINCIPAL RAY

Young's formulae (see, for example, Conrady 1929, p. 407) for computing astigmatism when the principal ray has finite angles of incidence are

$$\Delta\left\{N\left(\frac{\cos^2 I}{t}-\frac{\cos I}{r}\right)\right\}=0, \qquad \ldots\ldots(6)$$

$$\Delta\left\{N\left(\frac{1}{s}-\frac{\cos I}{r}\right)\right\}=0, \qquad \ldots\ldots(7)$$

where s and t are the distances along the principal ray to the sagittal and meridian foci and I is the angle of incidence. Each of these formulae must be computed with the full precision of an ordinary paraxial ray-trace, say five or six significant figures, in order to obtain the astigmatic difference at the end of the calculation with sufficient accuracy. The following method removes the necessity for this precision with equation (7).

The equation of a wave front given in equation (2) of paper I, can be written in the form

$$x=\frac{1}{2s}(y^2+z^2)+C_{20}y^2+\mathrm{O}(p^3), \qquad \ldots\ldots(8)$$

where p is a measure of the aperture and

$$C_{20}=\frac{1}{2}\left(\frac{1}{t}-\frac{1}{s}\right), \qquad \ldots\ldots(9)$$

so that C_{20} is the astigmatism coefficient.

From (6) and (7)

$$\Delta\{NC_{20}\cos^2 I\}=\tfrac{1}{2}N^2\sin^2 I\Delta\left(\frac{1}{Ns}\right); \qquad \ldots\ldots(10)$$

also, since $s^{(+1)}=s'-D'$, $t^{(+1)}=t'-D'$,

$$C_{20}{}^{(+1)}=\left(\frac{s'}{s^{(+1)}}\right)^2\frac{1}{1-\dfrac{2D's'}{s^{(+1)}}C_{20}{}'}C_{20}{}'. \qquad \ldots\ldots(11)$$

Equations (10) and (11) are the refraction and transfer formulae for C_{20}, and they can be applied when s and s' have been obtained by the use of equation (6), with the precision usual in aberration computations, e.g. three or four significant figures.

§ 4. COMPUTING EQUATIONS

The refraction and transfer formulae tabulated in § 8 of paper I and equations (2) to (5) of this paper give directly the values of the aberration coefficients of a pencil in terms of the constructional data of the system and the lower order properties of the pencil; for example s- and t-intersection lengths and coma coefficients are involved in the fourth order equations. Since the coefficients uniquely characterize the structure of the pencil in the neighbourhood of the principal ray, these equations may be considered to be in their most fundamental forms. However, for computing purposes it is more convenient to use equations involving actual wave-front aberrations instead of aberration coefficients, and these equations will be derived in the sub-sections which follow.

(i) *Axial Aberrations*

In deriving the computing equations for the axial aberrations the notation of Conrady (1929 p. 38) for paraxial ray-tracing will be used, with the additional symbol $\alpha=u+i$ for the paraxial 'central angle'. If y is the incidence height, at the surface considered, of the paraxial ray from the axial object point, the primary and secondary wave-front aberrations to be computed are defined by

$$W_4=NC_4y^4 \qquad \ldots\ldots(12)$$

$$W_6=NC_6y^6 \qquad \ldots\ldots(13)$$

respectively. Then if equation (2) is multiplied by y^4 and equation (3) by N the following computing formulae for W_4 are obtained:

$$\Delta W_4=\tfrac{1}{8}y(Ni)^2\Delta\left(\frac{u}{N}\right), \qquad \ldots\ldots(14)$$

$$\Delta_{\text{tr}}W_4=0, \qquad \ldots\ldots(15)$$

where Δ_{tr} denotes the increment on transfer between surfaces.

Similarly, on multiplying equation (4) by y^6 and equation (5) by N the computing formulae for W_6 are seen to be

$$\Delta W_6 = \tfrac{5}{2}(Ni)^2 \Delta\left(\frac{W_4}{N^2}\right) - 3\alpha N i \Delta\left(\frac{W_4}{N}\right) + \alpha^2 \Delta W_4 + \tfrac{1}{16} y (Ni)^3 \Delta\left(\frac{u^2}{N^2}\right), \qquad \ldots\ldots(16)$$

$$\Delta_{\text{tr}} W_6 = \frac{8d'}{N' y' y^{(+1)}} W_4'^2. \qquad \ldots\ldots(17)$$

By means of these computing equations the successive increments of W_4 and W_6 on refraction and transfer at each surface can be added, to obtain the values at the last surface in the system; finally the value of W_6 at the exit pupil can be obtained by applying equation (17), and then the values of C_4 and C_6 at the pupil can be obtained from equations (12) and (13).

The significance of W_6 can be explained as follows:

Let S_0 denote the paraxial ray of which the data are used in the above computations and let S denote the 'marginal' ray which coincided with S_0 at the first surface. The values of W_6 found for each surface are the secondary wave-front aberrations at points on the wave front of which the distances from the axis are the incidence heights of S_0; these are not, however, equal to the incidence heights of S after the first surface so that these values of W_6 do not refer to the points in which S meets the wave fronts. It can easily be shown that the difference, although usually numerically small, is mathematically of the order of magnitude of W_6, so that the values of W_6 do not refer to the ray S; they are, in fact, the aberrations at the incidence heights of S_0. Thus W_6 is primarily an auxiliary quantity which is used in computing the values of C_6, the aberration coefficient.

(ii) *Astigmatism*

In defining the wave-front aberrations which are used as auxiliary computing functions for the oblique aberration coefficients, certain convergence angles, etc. analogous to the paraxial u, i, α, y are used. These variables were defined by Hopkins (1946), who used them to give computing formulae, with checks, for tracing *s*- and *t*-fans along a principal ray. In this paper Hopkins' notation will be used with the addition of subscripts s or t to denote quantities referring to the sagittal or tangential sections respectively.

Formulae giving Hopkins' variables in terms of the results of an astigmatism calculation by means of Young's formulae can easily be obtained, but it is preferable either to calculate both the *s*- and *t*-traces directly by Hopkins' method or to calculate the *s*-trace only by Hopkins' method and then apply equations (10) and (11) to determine the astigmatism coefficient C_{20}; Hopkins' variables for the *t*-section can then be obtained by the following formulae:

$$\left.\begin{aligned} y_t' &= y_t \frac{\cos I'}{\cos I}, \\ u_t' &= \frac{y_t'}{s'} + 2y_t' C_{20}', \\ y_t^{(+1)} &= y_t' - D' u_t'. \end{aligned}\right\} \qquad \ldots\ldots(18)$$

The latter method has the advantage that part of the calculation (equations (10) and (11)) is carried out with quantities of aberrational magnitude.

(iii) *Third Degree Oblique Aberrations*

It is convenient to define the following symbols for use in this section and the next, all denoting quantities invariant on refraction:

$$\left.\begin{aligned} A&=Ni_{\mathrm{t}}\cos I, & A_{\mathrm{s}}&=Ni_{\mathrm{s}}, \\ h&=\frac{y_{\mathrm{t}}}{\cos I}, & G&=\frac{y_{\mathrm{s}}}{h}, \\ B&=N\sin I. \end{aligned}\right\} \qquad \ldots\ldots(19)$$

A quantity which refers to an aberration of the *n*th degree, i.e. varying as the *n*th power of the aperture, will have a subscript *ij*, where $i+j=n$; also a bracketed superscript (*m*) will denote a quantity referring to the *m*th refracting surface. Then the third degree aberrations which are used as computing functions are W_{30} and W_{12}, defined by

$$W_{30}=NC_{30}y_{\mathrm{t}}^3, \qquad \ldots\ldots(20)$$

$$W_{12}=NC_{12}y_{\mathrm{t}}y_{\mathrm{s}}^2, \qquad \ldots\ldots(21)$$

where C_{30} and C_{12} are aberration coefficients as defined in I. If equations (13) and (28) of paper I are multiplied by $(y_{\mathrm{t}}/\cos I)^3$ and $y_{\mathrm{t}}y_{\mathrm{s}}^2/\cos I$ respectively they become

$$\Delta W_{30}=\tfrac{1}{2}BhA\Delta\left(\frac{u_{\mathrm{t}}}{N\cos I}\right), \qquad \ldots\ldots(22)$$

$$\Delta W_{12}=\tfrac{1}{2}B\frac{y_{\mathrm{s}}^2}{r}\Delta(u_{\mathrm{t}})-\tfrac{1}{2}Bh\Delta(u_{\mathrm{s}}^2). \qquad \ldots\ldots(23)$$

By equations (36) and (37) of paper I $\Delta_{\mathrm{tr}}W_{30}=0$ and $\Delta_{\mathrm{tr}}W_{12}=0$, so that equations (22) and (23) are the computing formulae for W_{30} and W_{12}; if the system contains, say, *k* surfaces, the increments for each surface obtained from these equations are added to give $W_{30}'^{(k)}$ and $W_{12}'^{(k)}$; $C_{30}'^{(k)}$ and $C_{12}'^{(k)}$ can be obtained from these by means of equations (20) and (21), when they will refer to the wave front at the point at which the principal ray emerges from the last surface. Alternatively, the values of C_{30} and C_{12} can be made to refer to the wave front at the point where the principal ray in the image space crosses the axis, by using the transfer formulae of I.

(iv) *Fourth Degree Oblique Aberrations*

The wave-front aberrations used for the computing equations are defined by

$$W_{40}=NC_{40}y_{\mathrm{t}}^4, \qquad \ldots\ldots(24)$$

$$W_{22}=NC_{22}y_{\mathrm{t}}^2y_{\mathrm{s}}^2, \qquad \ldots\ldots(25)$$

$$W_{04}=NC_{04}y_{\mathrm{s}}^4. \qquad \ldots\ldots(26)$$

The procedure for deducing the equations is similar to that outlined in (iii), and the results are as follows:

Refraction equation for W_{40}:

$$\begin{aligned}\Delta W_{40}=\tfrac{1}{8}hA^2\Delta\left(\frac{u_{\mathrm{t}}}{N\cos I}\right)&+B\left[3A\Delta\left(\frac{W_{30}}{N^2\cos^2 I}\right)-\tfrac{3}{2}\alpha_{\mathrm{t}}\Delta\left(\frac{W_{30}}{N\cos I}\right)\right] \\ &+B^2\left[\tfrac{1}{2}hA\Delta\left(\frac{u_{\mathrm{t}}^2}{N^2\cos^2 I}\right)-\tfrac{1}{8}h\alpha_{\mathrm{t}}^2\Delta\left(\frac{u_{\mathrm{t}}}{N\cos I}\right)\right]. \qquad \ldots\ldots(27)\end{aligned}$$

Transfer equation for W_{40}:

$$\Delta_{\mathrm{tr}} W_{40} = \frac{9}{2} \frac{D'}{N' y_{\mathrm{t}}' y_{\mathrm{t}}^{(+1)}} W_{30}'^{2}. \qquad \ldots\ldots(28)$$

Refraction equation for W_{22}:

$$\begin{aligned}\Delta W_{22} = {} & \tfrac{1}{4} A \frac{y_{\mathrm{s}}^{2}}{r} \Delta(u_{\mathrm{t}}) - \tfrac{1}{4} h A \Delta(u_{\mathrm{s}}^{2}) + B \left[\frac{3}{2} \frac{y_{\mathrm{s}} G}{r} \Delta\left(\frac{W_{30}}{N \cos I}\right) + A \Delta\left(\frac{W_{12}}{N^{2} \cos^{2} I}\right) \right. \\ & \left. - \tfrac{1}{2} \alpha_{\mathrm{t}} \Delta\left(\frac{W_{12}}{N \cos I}\right) - 2 \alpha_{\mathrm{t}} \Delta\left(\frac{W_{12} \cos I}{N}\right) + 2 \frac{A_{\mathrm{s}}}{G} \Delta\left(\frac{W_{12}}{N^{2}}\right) \right] \\ & + B^{2} \left[\frac{1}{4} \frac{y_{\mathrm{s}}^{2} \alpha_{\mathrm{t}}}{r} \Delta\left(\frac{u_{\mathrm{t}}}{N \cos I}\right) - \tfrac{1}{2} A \frac{y_{\mathrm{s}}^{2}}{r} \Delta\left(\frac{u_{\mathrm{t}}}{N^{2} \cos^{2} I}\right) - \frac{1}{2} \frac{h}{G} \Delta\left(\frac{u_{\mathrm{s}}^{3}}{N}\right) \right].\end{aligned} \qquad \ldots\ldots(29)$$

Transfer equation for W_{22}:

$$\Delta_{\mathrm{tr}} W_{22} = \frac{3D'}{N' y_{\mathrm{t}}' y_{\mathrm{t}}^{(+1)}} W_{30}' W_{12}' + \frac{2D'}{N' y_{\mathrm{s}}' y_{\mathrm{s}}^{(+1)}} W_{12}'^{2}. \qquad \ldots\ldots(30)$$

Refraction equation for W_{04}:

$$\Delta W_{04} = \tfrac{1}{8} y_{\mathrm{s}} A_{\mathrm{s}}^{2} \Delta\left(\frac{u_{\mathrm{s}}}{N}\right) + \frac{1}{2} \frac{B y_{\mathrm{s}} G}{r} \Delta\left(\frac{W_{12}}{N \cos I}\right) + \tfrac{1}{8} B^{2} \frac{y_{\mathrm{s}}^{3}}{r^{2}} \left[\Delta\left(\frac{u_{\mathrm{s}}}{N}\right) - G \Delta\left(\frac{u_{\mathrm{t}}}{N \cos I}\right) \right]. \qquad \ldots\ldots(31)$$

Transfer equation for W_{04}:

$$\Delta_{\mathrm{tr}} W_{04} = \frac{1}{2} \frac{D'}{N' y_{\mathrm{t}}' y_{\mathrm{t}}^{(+1)}} W_{12}'^{2}. \qquad \ldots\ldots(32)$$

The successive increments for refraction and transfer at each surface are added to give the aberrations after refraction at the last surface; if desired the transfer equations can be used to find the values of W_{40}, W_{22} and W_{04} at the point where the principal ray crosses the axis in the image space; the values of the aberration coefficients for this point can then be found from equations (24) to (26), using appropriate values of y_{t} and y_{s}.

It should again be noted that, as in the case of W_6, the quantities $W_{40} \ldots W_{04}$ are not the aberrations of their respective types for one ray of the pencil throughout the system; they are the aberrations at points on the wave fronts of which the rectangular coordinates are $(x, y_{\mathrm{t}}, y_{\mathrm{s}})$ in the coordinate system used in I, and the differences between the aberrations at these points and those at the intersection points of an actual ray are not negligible. Thus W_{40}, W_{22} and W_{04} are auxiliary quantities, used for convenience in computing the aberration coefficients.

(v) *Field Lenses*

In field lenses the image is formed close to or on one surface, which for convenience may be called a field surface. As the image approaches a field surface, aberration coefficients tend to infinity, but it can be seen that the third degree aberrations W_{30} and W_{12} remain finite, since the infinite aberration coefficients refer to a pencil of zero aperture.

However, in certain cases, namely the transfer terms for W_{40}, W_{22} and W_{04}, and some terms in the refraction formulae for W_{22} and W_{04}, infinite values still occur, causing an indeterminacy on refraction at a surface where $t = 0$ or $s = 0$.

Dr. H. H. Hopkins suggested to the author that this indeterminacy might be removed by the device of computing the aberrations of a wave front at an infinite distance along the principal ray rather than at the refracting surface, and it was found that this could be carried out in the following manner: The aberrations at infinity are obtained from those at a refracting surface by putting $D=\infty$ in the transfer formulae, and they are then computed without using transfer formulae but with certain terms added to the refraction formulae; these additional terms transfer the incident wave front from infinity to the surface and transfer the refracted wave front back to infinity, and, together with those terms of the refraction formulae which become infinite, can be put into forms which are not indeterminate when $t=0$ or $s=0$. It is hoped to publish these results in detail at a later date.

(vi) *Summary of Computing Method*

To assess the state of correction of an optical system by the methods described would involve the following steps:

A paraxial ray from the axial object point is traced through the system and the data from this ray-trace are used in equations (14), (16), and (17) to determine the axial primary and secondary spherical aberration.

Several principal rays at different obliquities are traced trigonometrically or algebraically, and *s*- and *t*-fans are traced along these according to § 4 (ii) to obtain the data there referred to; these data are then substituted in equations (22) and (23) to obtain the third degree aberrations, and the results of this calculation together with the *s*- and *t*-trace data are used in equations (27) to (32) to obtain the fourth degree aberrations. The final aberration coefficients are obtained as described in the respective sections.

When all the aberration coefficients have been obtained for the points where the principal rays cut the axis they may be considered in that form or they may be combined as in equation (3) of paper I, to give the aberration coefficients in polar coordinates. The latter are, perhaps more useful when the performance in all azimuths, not merely meridian and sagittal, is being considered.*

When writing down the final expression for the wave-front aberration, whether in polar or rectangular coordinates, it usually happens (if the astigmatism is reasonably low) that the extra terms in t and s which occur are negligible compared with the main terms in the aberration coefficients.†

Hence for most purposes one can write for the aberration in rectangular coordinates

$$W=N\left\{\frac{1}{2}\left(\frac{1}{t}-\frac{1}{R}\right)y^2+\frac{1}{2}\left(\frac{1}{s}-\frac{1}{R}\right)z^2\right.$$
$$\left.+C_{30}y^3+C_{12}yz^2+C_{40}y^4+C_{22}y^2z^2+C_{04}z^4+\mathrm{O}(\rho^5)\right\}, \qquad \ldots\ldots(33)$$

---

* It seems to be impossible to obtain direct refraction formulae for the polar aberration coefficients, as may be seen, for example, by attempting to combine (13) and (28) of paper I according to (3) of paper I to obtain a refraction formula for B_{33}; it will be found that the angle of incidence cannot be eliminated from the left-hand side of the equation.

† In the example given in § 5 below the neglect of these terms caused an error of less than 0·1 wavelength at any point in the pupil for the pencil at 15° field angle; the error was less than 0·02 wavelength for the pencil at 12·5° field angle. The astigmatic differences in these two cases were about 0·7 inch and 0·2 inch respectively.

and in polar coordinates

$$W=N\left\{\frac{1}{2}\left(\frac{1}{t}-\frac{1}{s}\right)\rho^2\cos^2\phi+\frac{1}{2}\left(\frac{1}{s}-\frac{1}{R}\right)\rho^2+(C_{30}-C_{12})\rho^3\cos^3\phi\right.$$
$$+C_{12}\rho^3\cos\phi+(C_{40}-C_{22}+C_{04})\rho^4\cos^4\phi$$
$$\left.+(C_{22}-2C_{04})\rho^4\cos^2\phi+C_{04}\rho^4+\mathrm{O}(\rho^5)\right\}, \qquad \text{......(34)}$$

where R is the radius of the reference sphere, i.e. the distance along the principal ray from the wave front to the plane of focus considered. The values of the C's may refer either to the point where the principal ray leaves the last surface or to the point where it cuts the axis, and t and s should be chosen accordingly.

Four significant figures are sufficient in the aberration calculations, and in practice it often happens that only three or two figures of a number of terms are significant. In particular the contributions of the transfer terms to the fourth degree coefficients are often very small unless there is much uncorrected third degree aberration in a long air space.

The above procedure gives all the aberrations which affect the imagery of off-axis object points except distortion, which is obviously obtained directly from the principal ray-trace, and field curvature. Field curvature might be defined as the part of the wave-front aberration varying as $\rho^2 f(\sigma^2)$ (where $f(\sigma^2)$ is an even function of σ, the field angle) which is independent of astigmatism; this would define a generalized Petzval surface which would be the image surface in the absence of astigmatism. Although this definition is valid for primary aberrations, it is questionable whether it has a meaning when fourth and higher powers of σ are considered, since it is not obvious that an expansion for astigmatism involving σ^4 etc. would contain terms independent of the angles of incidence of the principal ray. However, the shapes of the s and t focal surfaces are more important for assessing the quality of the image, and these are known from the calculations. The choice of the radius of the reference sphere then enables one to consider the imagery on any focal surface, plane or otherwise.

§ 5. CLOSENESS OF APPROXIMATION OF THE FORMULAE

The formulae give the terms up to and including those of the fourth degree in the aperture in the expansion of the wave-front aberration as a power series, and it is desirable to determine the errors which occur through neglecting the remaining terms. The approach chosen was to investigate numerically systems of the types for which such a computing scheme is likely to be useful—systems of moderate aperture with fairly large angles of incidence of the principal ray. Several meridian and skew rays of the pencils surrounding the principal rays were traced through such a system and the resulting transverse aberrations were plotted and integrated graphically to determine the *true* wave-front aberrations; these were compared with the power series results. Some of the results of this comparison and a brief description of the method will be given here.

The work involved much skew ray tracing, and after several trials it was decided that the most convenient skew tracing method to use would be that described by Smith (1923); this method requires no trigonometrical tables, has no special cases where significant figures are lost, and can be checked completely at every stage.

To obtain the numerical integration formulae, let the equation of a wave front be equation (2) of paper I, and let (x, ρ, ϕ) be cylindrical polar coordinates of the point where a ray meets the wave front, so that $y=\rho\cos\phi$, $z=\rho\sin\phi$. Let (R, Y, Z) be the rectangular coordinates of the intersection of the ray with an image plane whose equation is $x=R$. Then by geometry

$$Y=\rho\cos\phi-\left\{\frac{\partial x}{\partial\rho}\cos\phi-\frac{1}{\rho}\frac{\partial x}{\partial\phi}\sin\phi\right\}(R-x)+\mathrm{O}(\rho^4),$$

$$Z=\rho\sin\phi-\left\{\frac{\partial x}{\partial\rho}\sin\phi+\frac{1}{\rho}\frac{\partial x}{\partial\phi}\cos\phi\right\}(R-x)+\mathrm{O}(\rho^4),$$

so that

$$Y\cos\phi+Z\sin\phi=\rho-(R-x)\frac{\partial x}{\partial\rho}+\mathrm{O}(\rho^4).$$

On integrating this with respect to ρ, substituting the value of x in polar coordinates, and re-grouping the terms, one obtains

$$\begin{aligned}
-\frac{1}{R}\int_0^\rho\{Y\cos\phi+Z\sin\phi\}\,d\rho+\frac{1}{8R}\left\{\left(\frac{1}{t}-\frac{1}{s}\right)\cos^2\phi+\frac{1}{s}\right\}^2\rho^4-\frac{1}{8R^3}\rho^4\\
=\frac{1}{2}\left(\frac{1}{t}-\frac{1}{s}\right)\rho^2\cos^2\phi+\frac{1}{2}\left(\frac{1}{s}-\frac{1}{R}\right)\rho^2+(C_{30}-C_{12})\rho^3\cos^3\phi+C_{12}\rho^3\cos\phi\\
+\left\{C_{40}-C_{22}+C_{04}+\frac{1}{8}\left(\frac{1}{t^3}-\frac{2}{t^2s}+\frac{1}{s^3}\right)\right\}\rho^4\cos^4\phi\\
+\left\{C_{22}-2C_{04}+\frac{1}{4s}\left(\frac{1}{t^2}-\frac{1}{s^2}\right)\right\}\rho^4\cos^2\phi+\left\{C_{04}+\frac{1}{8}\left(\frac{1}{s^3}-\frac{1}{R^3}\right)\right\}\rho^4+\mathrm{O}(\rho^5).
\end{aligned}\qquad\ldots\ldots(35)$$

This equation is arranged so that the right-hand side gives the wave-front aberration terms up to those in ρ^4, plus the remainder term, while the left-hand side gives the *true* wave-front aberration of all orders; the first term on the left-hand side is the integral of the transverse aberration referred to above. In practice several terms in equation (35) are usually negligible, as explained at the end of § 4, and it then takes the simpler form

$$\begin{aligned}
-\frac{1}{R}\int_0^\rho(Y\cos\phi+Z\sin\phi)\,d\rho=\frac{1}{2}\left(\frac{1}{t}-\frac{1}{s}\right)\rho^2\cos^2\phi+\frac{1}{2}\left(\frac{1}{s}-\frac{1}{R}\right)\rho^2\\
+(C_{30}-C_{12})\rho^3\cos^3\phi+C_{12}\rho^3\cos\phi+(C_{40}-C_{22}+C_{04})\rho^4\cos^4\phi\\
+(C_{22}-2C_{04})\rho^4\cos^2\phi+C_{04}\rho^4+\mathrm{O}(\rho^5).
\end{aligned}\qquad\ldots\ldots(36)$$

To use this equation three or four rays are traced at each value of ϕ to be considered, Y and Z are determined for each ray, and the integral on the left-hand side is evaluated graphically and tabulated as a function of ρ. This is compared with the right-hand side, also considered as a function of ρ, and the difference gives the remainder terms of order ρ^5 and higher which should occur in the power series for the wave-front aberration. In the examples to be given, the axial secondary spherical aberration was added to the right-hand side, to improve the fit.

The results to be given here were from calculations on a design of a 20-inch F/6·3 triplet objective of which the specification was kindly made available by Messrs Kodak Ltd. The objective was designed to work up to 10° semi-field angle, but the calculations were made for angles up to 15°, since the aberrations were still of a suitable magnitude for this purpose.

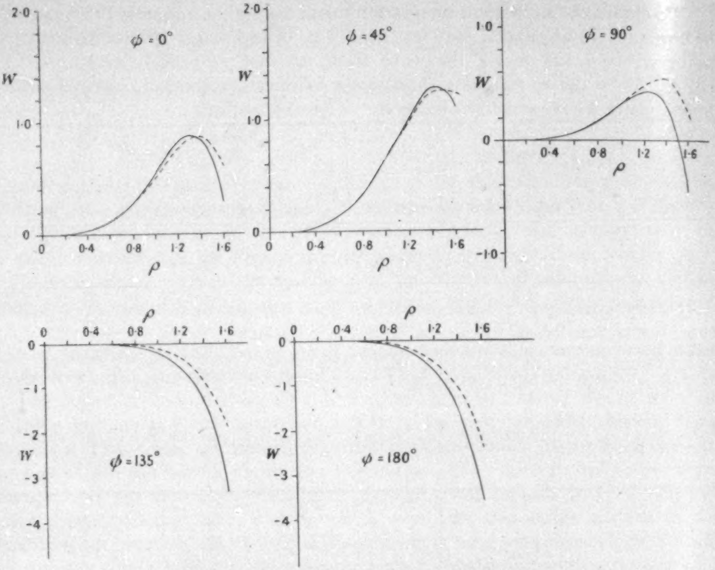


Figure 1. Wave-front aberration at $U_{pr} = -70°$.
Aberration W in microns; semi-aperture ρ in inches.
Full lines : from ray tracing. Dotted lines : from aberration coefficients.

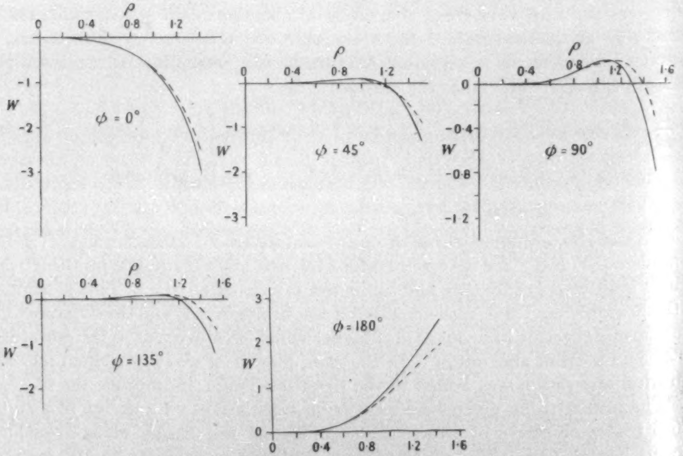


Figure 2. Wave-front aberration at $U_{pr} = -10°$.
Aberration W in microns; semi-aperture ρ in inches.
Full lines : from ray tracing. Dotted lines : from aberration coefficients

The aberrations are given in microns of optical path and the aberration coefficients are in microns per inch^3 and microns per inch^4 for the third and fourth degree coefficients respectively. The coefficients for rectangular coordinates, referred to the exit pupil, are tabulated against U_{pr}', the convergence angle of the emerging principal ray, and these can be directly substituted in equations (33)

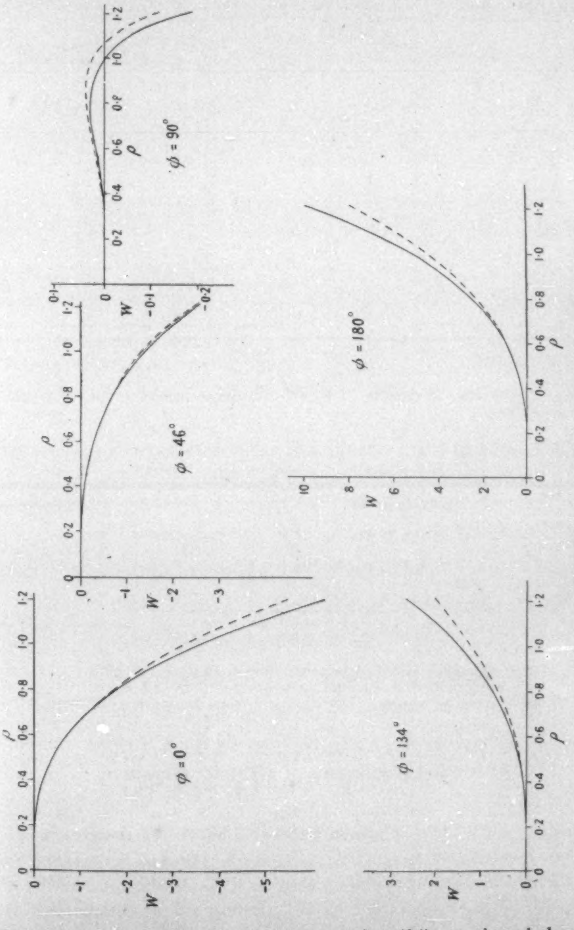


Figure 3. Wave-front aberration at $U_{pr} = -12\frac{1}{2}°$.
Aberration W in microns, semi-aperture ρ in inches.
Full lines : from ray tracing. Dotted lines : from aberration coefficients.

and (34); the results of substituting in equation (34) are given below the table. In the graphs W, the aberration as found from the coefficients (broken lines) and by ray-tracing (full lines) is plotted against ρ, the semi-aperture, for various azimuths at each value of U_{pr} (Figures 1 to 4), and for an axial pencil (Figure 8). The reference sphere in each case is chosen to give no second degree aberration (error of focus) in the azimuth concerned. The *s*- and *t*-fields and the polar aberration coefficients are plotted against U_{pr} in Figures 5, 6 and 7.

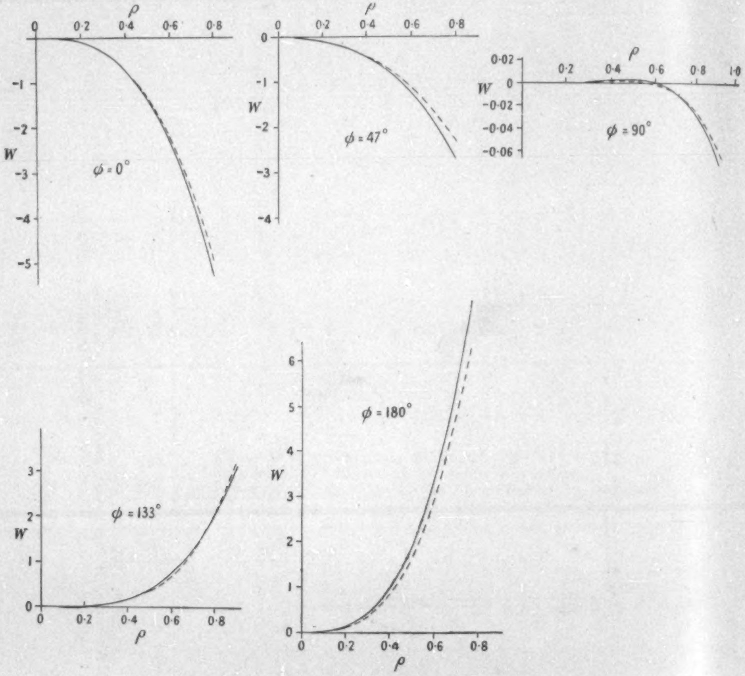


Figure 4. Wave-front aberration at $U_{pr}=-15°$.
Aberration W in microns, semi-aperture ρ in inches.
Full lines : from ray tracing. Dotted lines : from aberration coefficients.

Aberration Coefficients of a Triplet Objective

U_{pr}	U_{pr}'	C_{12}	C_{30}	C_{40}	C_{22}	C_{04}
(degrees)		(microns/inch^3)		(microns/inch^4)		
0	0	0	0	0·649	1·298	0·649
− 7	− 7·0237	0·713	0·350	0·325	0·856	0·497
−10	−10·0406	0·658	− 0·778	0·272	0·425	0·349
−12·5	−12·5655	0·056	− 3·85	0·858	−0·150	0·201
−15	−15·1126	−1·700	−11·95	2·77	3·33	0·028

Axial secondary spherical aberration : $C_6=-0·177$ microns/inch^6.

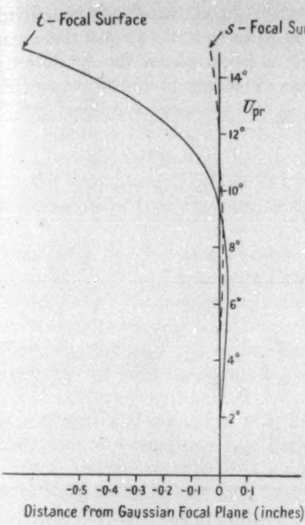


Figure 5. *s*- and *t*-focal surfaces.

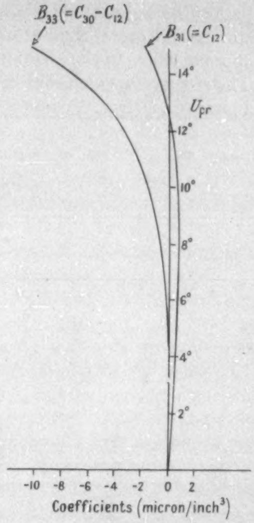


Figure 6. Third degree polar aberration coefficients.

Wave-front aberration in polar co-ordinates (from equation (34)), with reference sphere centred on the sagittal focus:

$$\begin{array}{cl}
U_{\mathrm{pr}}\ \text{(degrees)} & \qquad\qquad W\ \text{(microns)} \\
-7 & \left\{\begin{array}{l} -0{\cdot}808\,\rho^2\cos^2\phi-0{\cdot}363\,\rho^3\cos^3\phi+0{\cdot}713\,\rho^3\cos\phi-0{\cdot}034\,\rho^4\cos^4\phi \\ -0{\cdot}138\,\rho^4\cos^2\phi+0{\cdot}497\,\rho^4-0{\cdot}177\rho^6 \end{array}\right. \\
-10 & \left\{\begin{array}{l} 0{\cdot}615\,\rho^2\cos^2\phi-1{\cdot}436\,\rho^3\cos^3\phi+0{\cdot}658\,\rho^3\cos\phi+0{\cdot}196\,\rho^4\cos^4\phi \\ -0{\cdot}273\,\rho^4\cos^2\phi+0{\cdot}349\,\rho^4-0{\cdot}177\,\rho^6 \end{array}\right. \\
-12{\cdot}5 & \left\{\begin{array}{l} 5{\cdot}95\,\rho^2\cos^2\phi-3{\cdot}91\,\rho^3\cos^3\phi+0{\cdot}056\,\rho^3\cos\phi+1{\cdot}209\,\rho^4\cos^4\phi \\ -0{\cdot}552\,\rho^4\cos^2\phi+0{\cdot}201\,\rho^4-0{\cdot}177\rho^6 \end{array}\right. \\
-15 & \left\{\begin{array}{l} 20{\cdot}3\,\rho^2\cos^2\phi-10{\cdot}25\,\rho^3\cos^3\phi-1{\cdot}70\,\rho^3\cos\phi-0{\cdot}53\,\rho^4\cos^4\phi \\ +3{\cdot}39\,\rho^4\cos^2\phi+0{\cdot}03\,\rho^4-0{\cdot}177\,\rho^6 \end{array}\right.
\end{array}$$

§ 6. DISCUSSION OF THE RESULTS

The purpose of analytical aberration calculations is to obtain a useful approximation to the aberrations, with an indication of the relative contributions of the various components of the system. It is immaterial in most cases if the *final* result is in error by say, 10%, since tolerances cannot be assigned with such precision. An exact knowledge of the aberrations, such as might be obtained by extensive ray tracing, would add nothing of further value for assessing the general performance and would give no indication as to the contributions of individual surfaces. From this point of view it can be seen from the graphs that a useful approximation to the wave-front aberration is obtained for the aperture used. The angles of incidence of the principal rays ranged up to 26° on the air side of a glass–air surface.

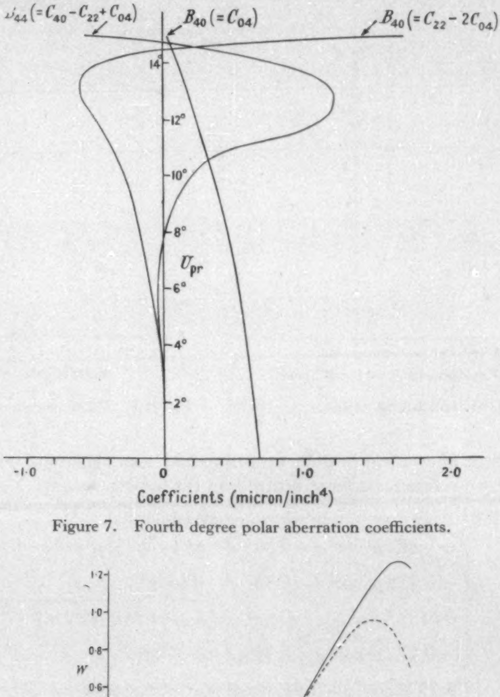


Figure 7. Fourth degree polar aberration coefficients.

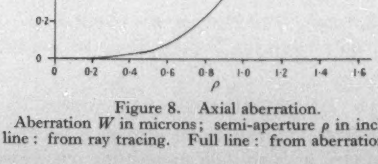


Figure 8. Axial aberration.
Aberration W in microns; semi-aperture ρ in inches.
Dotted line: from ray tracing. Full line: from aberration coefficients.

Some interesting points emerge from a study of the graphs showing the variations of the coefficients with field angle. It can be seen that C_{30} is zero when U_{pr} is $-8\frac{1}{2}°$, so that from meridian ray-traces one would draw the conclusion that there was no coma at this field angle; however, the value of C_{12} shows that at 45° azimuth in the pupil there is about $0{\cdot}5\,\mu$ of coma. Again, at $U_{pr} = -9°$

the value of C_{40} shows about $0{\cdot}2\,\mu$ of spherical aberration in the meridian plane, but there is $0{\cdot}4\,\mu$ in the sagittal plane and the run of the coefficients shows that it is almost $0{\cdot}4\,\mu$ at nearly all azimuths out of the meridian plane. On the other hand, at -15° there is about $3\,\mu$ of spherical aberration in the meridian plane, but this rapidly decreases to nearly zero in the sagittal plane.

In Figure 6 the shape of the B_{31} graph (coma varying in the aperture as $\rho^3 \cos\phi$) shows that in addition to the Seidel component, which varies linearly with the field angle U_{pr} there is a component of opposite sign varying as $U_{\mathrm{pr}}{}^3$ (an aberration of type $\rho^3\sigma^3 \cos\phi$ in the usual aberration classification mentioned in I).

Similarly the B_{33} graph (coma varying as $\rho^3 \cos^3\phi$) contains components varying as $U_{\mathrm{pr}}{}^3$, the lowest power which could occur, and as $U_{\mathrm{pr}}{}^5$.

In Figure 7 the ordinary spherical aberration coefficient (B_{40}, for aberration varying as ρ^4) shows, in addition to the constant axial component, a part varying as $U_{\mathrm{pr}}{}^2$; B_{42}, the coefficient of the aberration varying as $\rho^4 \cos^2\phi$, shows components varying as $U_{\mathrm{pr}}{}^2$ and $U_{\mathrm{pr}}{}^4$, while B_{44} has components varying as $U_{\mathrm{pr}}{}^4$ and $U_{\mathrm{pr}}{}^6$.

The general conclusion may be drawn that for systems of moderate aperture and extended field the use of this method of assessment will show up properties which could not otherwise have been discovered without a large amount of ray-tracing. An estimate based on the time taken for the above computations suggests that to determine the final wave-front aberrations by ray-tracing would take about four times as long as the calculation of the aberration coefficients. An additional advantage is that the aberration calculations give immediately the contributions of each part of the system to the final aberrations.

ACKNOWLEDGMENTS

This work is published by permission of the Chief Scientist, the Ministry of Supply. Thanks are due to Professor L. C. Martin and Dr. H. H. Hopkins for advice and criticism, and to Miss J. M. Drewitt, who did the computations.

REFERENCES

Conrady, A. E., 1929, *Applied Optics and Optical Design*, Part I (Oxford : University Press).
Hopkins, H. H., 1946, *Proc. Phys. Soc.*, **58**, 663.
von Rohr, M., 1920, *The Formation of Images in Optical Instruments*, translated by R. Kanthack (London : His Majesty's Stationery Office).
Smith, T., 1923, *Dictionary of Applied Physics*, Vol. IV (London : Macmillan and Co. Ltd.), article on " Optical Calculations ".
Weinstein, W., 1949, *Proc. Phys. Soc.* B, **62**, 726.

LETTERS TO THE EDITOR

An Optical Method for Studying the Deformation of Mild Steel

The purpose of this note is to indicate some of the work which has continued over the past few years at the Cavendish Laboratory on the Lüders' deformation of mild steel. In particular, a new ' optical strain gauge ' method will be described, which has been used in studying the properties of strain-aged materials.

When a normal tensile test-piece of mild steel is strained, the stress–strain curve which results is very jagged, because of the generation of a complex system of Lüders' bands. Some years back, however, Dr. E. Orowan and Mr. W. Chitruk found that this difficulty could be avoided by the use of thin wire specimens. The Lüders' band is nucleated at one grip, after a marked upper yield point, and spreads over the length of the specimen under virtually constant stress. The investigation was carried a stage further here by Mr. Sylwestrowicz in work for the British Iron and Steel Research Association. He made a preliminary study of the yield point and strain-ageing properties of the material, and found that, if a Lüders' band is propagated along a specimen, the load released, and the specimen aged for short times in the temperature interval 50° c.–200° c., a new upper yield point appears before the band spreads further. After longer ageing times, however, the front of the band is so securely locked that a new band appears from the grip, propagating in the *old* one. This new band will be called the ' secondary ' Lüders' band, and the original the ' primary ' band. The propagation stress for the secondary band may be anything up to 40% more than the primary one.

In the course of a systematic study by the writer of the strain-ageing properties of these wire specimens, it was necessary to devise some method which would allow the development of the secondary band to be followed. A technique has been adopted which was originally used in a study of the jerky extension of zinc crystals by Orowan (1949). Fine filings of tin are dropped down a tube of molten stearic acid. The heating arrangements are such that the temperature at the top of the tube is above the melting point of tin (232° c.) and below it at the bottom. Thus, the tin filings melt, and solidify as balls at the bottom. The molten stearic acid is then poured off and, when cool, the balls are extracted and cleaned with ether.

The balls are mounted half-way down the wire tensile specimen on a thin thread of rubber solution, and are illuminated by a Western ' concentrated arc ' lamp. The images of the source in the balls are focused on to a strip of photographic paper in a drum camera. A good lens of short focal length must be used, giving an optical magnification of × 15 to × 50.

Typical records are shown diagrammatically in Figure 1. In the extensometer, the bottom end of the specimen is pulled down, and the consequent traces on the record are the inverted pictures of the motion of each ball. The start of each trace is under zero stress. As the load comes on, the balls move downwards with a velocity of $V/2$. When a band is nucleated at the bottom grip, as shown in Figure 1 (*a*), no deformation occurs outside the band, the balls remaining fixed in space, and the trace is horizontal. As the edge of the band moves past the balls, all the deformation occurs above them, and they move downward with the full velocity of the straining head. Thus, the gradient of the trace is twice that in the elastic region. The line of transition is staggered across the record, as shown by the hypotenuse of the dotted triangle in the diagram. When the band has spread to the top grip, deformation is from then on homogeneous, and the gradient of the trace is the same as in the elastic region.

Figure 1 (*b*) shows the traces in the similar case when the band is nucleated in the top grip, while Figure 1 (*c*) shows the more rare example of two bands generated simultaneously.

Figure 2 (*a*) shows the photograph of a typical record, obtained from undeformed material, and similar in form to Figure 1 (*a*) above. From the gradients of the traces, and knowing the velocity V of the straining head and the rate of rotation of the camera drum, we may determine the optical magnification, the rate of advance of the band edge, and the Lüders' strain. For this particular record, the values obtained are × 18, 0·76 cm/min., and 4·3% respectively.

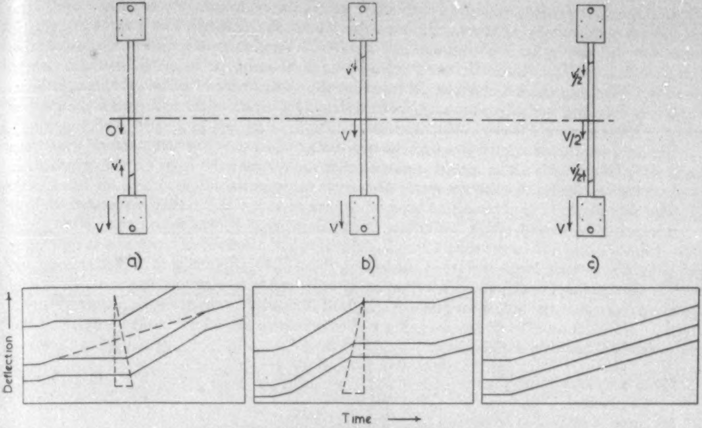

O
V'
V
V'
V
V
V'/2
V/2
V'/2
V
a)
b)
c)
Deflection
Time

Figure 1. Diagram of typical records.

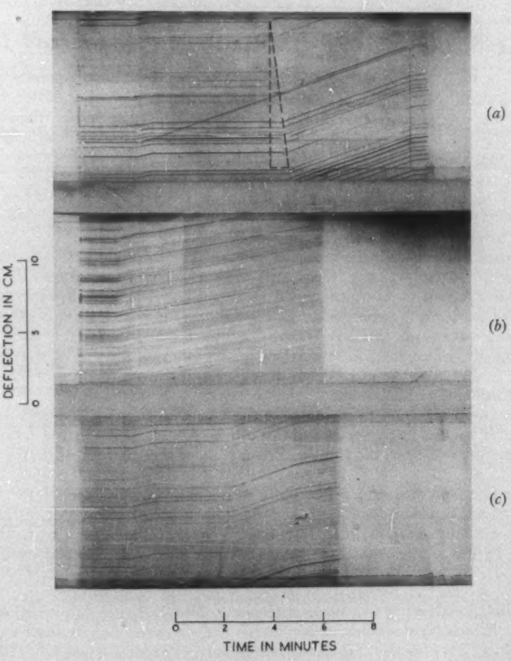


Figure 2. Photograph of records.

If the specimen is strained to the end of the Lüders' extension, and then aged, the secondary band slowly appears. Figure 2 (*b*) shows the deformation after $4\frac{1}{2}\%$ plastic extension and ageing for 30 minutes at 60° c. Here, a band has started from the lower grip, but material outside the band is deforming simultaneously, perhaps as localized bands. Soon the band edge meets material deformed to the same order of plastic strain as its own Lüders' strain, and so the band spreads no further. Deformation then spreads from another point. All this occurs under constant stress.

Figure 2 (*c*) shows a later stage of the secondary band development, after $4\frac{1}{2}\%$ strain and one hour at 60° c. The record is now similar to (*a*), but even here a little deformation occurs outside the band, and the traces are never quite horizontal.

The mechanics of this transition type of deformation are still under investigation.

One other important result has arisen from this work. It was found that if the grain size of the specimen is larger than a certain critical value, Lüders' deformation is no longer possible. For a basic Bessemer steel, containing 0·045% C, this value is 60 ± 3 grains/mm. For coarser grained specimens, deformation is still under constant stress, but the traces obtained show that the actual mechanism involved is similar to the transition type obtained above. No amount of straining or ageing will ever induce a Lüders' band to appear if the grain size is above this critical limit.

Cavendish Laboratory,
Cambridge.
5th June, 1950.

E. O. HALL.

OROWAN, E., 1949, *J. Metals* (*Metals Trans.*), **185**, 876

The Scattering of 3-cm. Radiation by Ionized Gases

Work has been carried out in this laboratory on ultra-high-frequency techniques for studying electron concentrations in discharges, particularly on the scattering of radiation from low pressure discharges (cf. the meteor work, especially the theoretical work of Herlofson (1948 a, b)) for wavelengths of 3 cm., but also on the absorption in discharges of such radiations (see Andrew *et al.* 1948, Adler 1949 for similar recent work on absorption). Since, so far as we know, no other published account of scattering from laboratory discharges has been given, although experiments on the aurora have been reported (Lovell *et al.* 1947, Forsyth *et al.* 1950), it was thought worth while to report the following preliminary experiments. The experiments on absorption will be reported elsewhere.

Experiments have been made with two sealed-off commercial mercury lamps of diameters approximately 3·1 cm. (17 cm. long) and 6·4 cm. (40 cm. long), and the experimental arrangement comprises transmitting and receiving horns arranged prependicular to each other so that scattered radiation from the discharge is measured, using a crystal matched to the receiving waveguide (VSWR 1·08). Figures 1 and 2 show the scattered power in arbitrary units as a function of the tube currents (note that the Hg vapour pressure varies, because of temperature variations, with the current) for two directions of the electric vector (perpendicular and parallel to the discharge axis).

The scattered radiation from the glass tube, at zero current, is particularly important in the case of Figure 1 and the possibility of interference between the radiation scattered from the plasma and from the front and rear walls of the tube for all cases is an obvious factor.

If the effects of magnetic fields, both strays and the self-field of the discharge, are neglected and if collision damping is also negligible, the electron concentration n may be found, if the scattering shows a sharply defined onset and is equivalent to the ionospheric refraction, by the simple critical frequency formula (dielectric constant $\epsilon=0$)

$$n_{\text{crit}}=1{\cdot}24\times10^{-8}f^2;$$

f is the radiation frequency in cycles per second. For $\lambda=3$ cm., $n_{\text{crit}}=10^{12}/\text{cm}^3$, and this is not unreasonable as a maximum axial electron concentration for the discharges considered (approximately 5% ionization) at the currents shown in Figures 1 and 2. This treatment is, no doubt, an over-simplification and the following comments are pertinent.

1. The three cases of radiation scattering (wavelength λ), for discharges of diameter d are: (*a*) extended scatterer of dimensions very much greater than d (ionosphere, see for example, Mitra 1949); (*b*) cylindrical discharge, $d \ll \lambda$ (meteor trails), where the scattering may be calculated as shown by Blackett and Lovell (1941), Lovell and Clegg (1948) and Lovell (1950) if ϵ is of the order unity; (*c*) cylindrical discharge, d, of the order of λ (present discharges) where the radiation, scattered across a diameter, shows phase differences (Herlofson 1948) if ϵ is of the order unity.

2. Herlofson states that when ϵ is approximately zero (present discharges) at the centre of the scattering column, the scattering with the electric vector parallel to the discharge column can be calculated as in cases (1 *b*) and (1 *c*) up to 10^{13} electrons per centimetre length of scatterer and may be much less than when the electric vector is perpendicular to the discharge column (perhaps by several powers of 10), since in the latter case plasma oscillations may be excited. This type of effect seems to be apparent with the small tube (Figure 1) but not with the large tube (Figure 2).

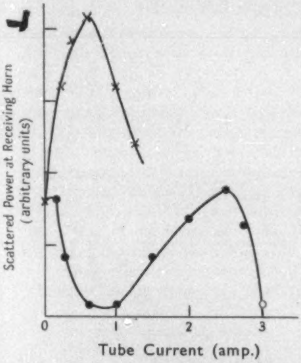


Figure 1. Small tube.

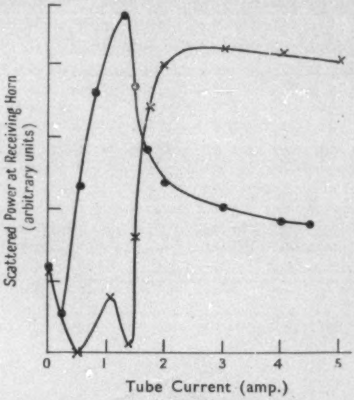


Figure 2. Large tube.

● *E* vector parallel to tube axis.
× *E* vector perpendicular to tube axis.

Units on the ordinate axis for the two curves are not comparable since the tube dimensions are different, but the curves are normalized at zero current.

The experiments are being extended in various directions and will be reported in more detail, with the absorption work, elsewhere. It is intended to compare the electron concentration measurements with those obtained with probe and spectroscopic techniques.

Department of Electrical Engineering,
The University of Liverpool.
16th April 1950, in final form 3rd July 1950.

S. N. DENNO.
H. A. PRIME.
J. D. CRAGGS.

ADLER, F. P., 1949, *J. Appl. Phys.*, **20,** 1125.
ANDREW, E. R., AXFORD, D. W. E., and SUGDEN, T. M., 1948, *Trans. Faraday Soc.*, **44,** 427.
BLACKETT, P. M. S., and LOVELL, A. C. B., 1941, *Proc. Roy. Soc.* A, **177,** 183.
FORSYTH, P. A., PETRIE, W., VAWTER, F., and CURRIE, B. W., 1950, *Nature, Lond.*, **165,** 561.
HERLOFSON, N., 1948 a, *Observatory*, **68,** 230; 1948 b, *Rep. Prog. Phys.*, **11,** 444 (London : Physical Society).
LOVELL, A. C. B., 1950, *Sci. Prog.*, **38,** 22, and references there cited.
LOVELL, A. C. B., and CLEGG, J. A., 1948, *Proc. Phys. Soc.*, **60,** 491.
LOVELL, A. C. B., CLEGG, J. A., and ELLYETT, C. D., 1947, *Nature, Lond.*, **160,** 372.
MITRA, S. K., 1949, *The Upper Atmosphere* (Calcutta).

Temperature Dependence of Counter Characteristics in Self-Quenching Geiger-Müller Counters.

Parkash and Kapur in a recent paper (1950) have reported on the variation of characteristics with temperature of the usual type of argon–alcohol counter. They used counters of two types, viz. those in internal copper cathodes and those with external graphite cathodes. They noted that in counters with internal cathodes, the plateau decreased in length and disappeared at about 9° c., but that no such effect was obtained with the second type of counter. At higher temperatures (up to 62° c.) they obtained an increase of plateau slope and a decrease of plateau length in both cases.

Our experience at Harwell concerning the effect of increase of temperature on plateau slope and plateau length confirms the results of Parkash and Kapur, but we do not agree with the deductions they make about the advantage of the external cathode counter (' Maze ' tube) at low temperatures.

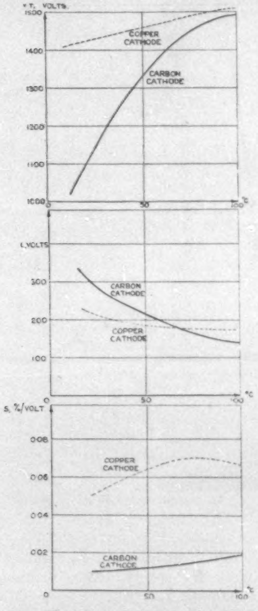


Figure 1. Variation of threshold voltage (V_T), plateau length (L) and plateau slope (S) with change of temperature for a counter with argon–alcohol filling.
——— internal metal cathode counter.
– – – – internal graphite cathode counter.

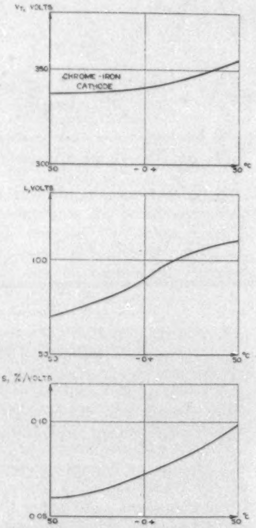


Figure 2. Variation of threshold voltage (V_T), plateau length (L) and plateau slope (S) with change of temperature for an internal metal cathode counter with an argon–neon–bromine filling gas.

Typical results, obtained at Harwell, showing the variation of threshold voltage V_T, plateau length L and plateau slope S with change of temperature are shown in Figure 1. Two types of counters were used for these tests, viz. (i) a counter with an internal metal cathode and (ii) a counter with an internal graphite cathode. The filling gas was the argon–alcohol mixture as used by Parkash and Kapur and the change of characteristics with increase of temperature is as reported by these authors. It will be noted that the temperature coefficient, defined as the average change of threshold voltage per degree Centigrade, is higher for the graphite cathode counter than for the metal cathode counter. This is to be expected, since the adsorption of alcohol vapour by a graphite surface is very temperature dependent.

Tests have also been made at Harwell using the Maze (1946) type of counter with an external cathode on the outside of a glass envelope. The temperature coefficient obtained was very low as compared with that obtained with the internal cathode type of counter, as also reported by Parkash and Kapur, but in our tests it was noted that the performance was markedly affected by the electrical resistance of the glass. In practice, the electrical resistance of the glass becomes too high at low temperatures to allow this type of counter to operate. For this reason we do not agree with the conclusion of Parkash and Kapur that the Maze type of counter with the usual argon–alcohol filling is satisfactory for low temperature operation. The Maze type of counter has several advantages, e.g. simple construction, great robustness (excess voltages may be used without destroying the characteristics), etc., but like all counters using ethyl alcohol as the quenching agent it suffers from the disadvantage that the alcohol condenses at low temperatures. In our view, if low temperature operation is required, it is better to use a counter with an internal cathode, but with a quenching agent other than ethyl alcohol. When operation at temperatures of the order of 5° C. or below has been a requirement, we have used ethyl bromide as the quenching agent, but more recently we have used bromine, and now all our Geiger counter instruments for geological prospecting (for which operation over a wide temperature range is required) are using counters with this quenching agent. Some typical results for a counter of this type are shown in Figure 2. Counters with bromine as the quenching agent have other advantages, as has already been reported (Franklin and Loosemore 1950, Taylor 1950). Of these other advantages, probably their long life ($\sim 10^{10}$ counts) is the most important.

Atomic Energy Research Establishment,
Harwell, Berkshire.
3rd July 1950.

W. R. Loosemore.
Denis Taylor.

Franklin, E., and Loosemore, W. R., 1950, *J. Instn. Elect. Engrs.*, in press.
Maze, R., 1946, *J. Phys. Radium*, **7**, 164
Parkash, O., and Kapur, P. L., 1950, *Proc. Phys. Soc.* B, **63**, 461.
Taylor, D., 1950, *J. Sci. Instrum.*, **27**, 81.

REVIEWS OF BOOKS

Some Recent Researches in Solar Physics, by F. HOYLE. Pp. xii + 134. Cambridge Monographs on Physics. 1st Edition. (Cambridge: University Press, 1949). 12*s*. 6*d*.

Mr. Hoyle has given us a revolutionary little book which will have a profoundly stimulating effect upon the development of solar physics.

Without prejudice to the splendid contributions of the pioneer workers of fifty years ago, such as Lockyer, Evershed and Hale, we may say that the present situation in solar physics is the outcome of the development of new experimental techniques developed within the last twenty years. The author has been quick to accept the challenge presented by this accumulation of new knowledge, and he has worked out in detail an entirely novel approach to the main problem. This is the *accretion hypothesis* which he has developed jointly with Lyttleton and Bondi.

The central problem may be briefly stated as follows. All our preconceived ideas would lead us to suppose that the sun's outer atmospheric layers are cooler than those further in. In fact, the reverse is the case. The concentration of the excited states of atoms of a given element can be calculated as a function of height in the solar chromosphere, assuming the temperature to be that of the reversing layer (4830° K.) at its base where the Fraunhofer spectrum is formed. Eclipse observations of the chromospheric spectrum, however, have shown that many lines, especially those of hydrogen and helium, extend to several times the predicted heights. A higher temperature in the chromosphere is evidently required, and a value of the order of 30,000° K. has been confirmed spectroscopically by Redman, who attributes the observed widths of the chromospheric hydrogen lines to thermal broadening.

Above the chromosphere is the corona. Its spectrum could not be more different. The emission lines, as studied by Lyot and Waldmeier and interpreted by Edlén, are those of highly ionized metals. Three independent experimental determinations agree upon a value for the kinetic temperature of the coronal material of the order of 10^6 deg. K. How then is this inverted temperature gradient maintained? Hoyle's reply is the bold, yet obvious, assertion that energy is being supplied to the top of the coronal atmosphere by the inflow of interstellar material under the action of gravitational attraction.

Accretion of interstellar material is no new idea; but the theory has never been worked out in all its implications and with such a wealth of quantitative detail as has been done by Mr. Hoyle and his collaborators. It is the main theme of the book, as set out in Chapters III and IV, but there is much more in addition. The author has a genius for sifting through the observational material and for selecting what are the most significant features in relation to his own special problems. Thus Chapters I and II are mainly concise summaries of the observational data concerning sunspots, the solar cycle, prominences, chromosphere and corona. In Chapter V he puts forward what promises to be a very fruitful theory of prominences, regarding them as cooler regions of the corona where the temperature is low enough for many of the normal chromospheric lines to become visible. There are new ideas on the sunspot magnetic cycle, and Giovanelli's work on solar flares is described and extended. Chapters V and VI deal with solar and terrestrial relationships and with the emission of radio waves from the sun.

In an appendix which discusses the origin of the sun's general magnetic field the theory is advanced that the sun's present magnetic field may have originated within the last million years or so during one of the temporary periods of very rapid accretion of interstellar matter. Indeed, it is one of the virtues, if we may call it such, of the accretion hypothesis that it allows a certain degree of flexibility, so to speak, in the ordering of solar affairs, and evidence for this is not wanting. As the sun wends its dusty way through interstellar space it may sweep up the interstellar particles from time to time at very different rates. During periods of high accretion rate solar activity will be on a grander scale than at present. Mr. Hoyle thinks that during one of these periods the earth may have received its present magnetization through streams of ions ejected from solar flares far more intense than those which now generate terrestrial magnetic storms.

Enough has been said to convey the impression that this is a book full of new and challenging ideas : many of them have been worked out in full detail, others are no more than daring speculations which may, or may not, survive. However that may be, it is certain that this little volume contains a sufficient amount of solid achievement to place solar physicists in Mr. Hoyle's debt for many years to come. M. A. ELLISON.

Atmospheric Turbulence, by O. G. SUTTON. Pp. viii + 107. Methuen's Monographs on Physical Subjects. First Edition. (London: Methuen, 1949). 6*s*.

Professor Sutton has been a leader during the past two decades in the development of the British school on atmospheric turbulence and he gives us here a well-written account of that development. To the reader a little outside the subject he provides an opportunity to see how far along the road of understanding some parts of the subject have progressed, and, what is clearly not the same thing, how far the phenomena dealt with have been made amenable to mathematical treatment.

This is a monograph and, fittingly enough, the author has followed a personal bent in dealing with his subject wherever possible by mathematical methods. But one should hasten to add that no physicist should find the matter difficult to read on that account. He further limits himself in the main to a study of the problems of the turbulent diffusion of momentum, heat and matter in the lowest 100 metres or so of the atmosphere; a quite understandable choice since it is in that layer that knowledge has in the main advanced in recent years. Nevertheless in making the choice the author claims to be dealing " with the most important region for turbulence " and " to have indicated to the reader those parts of the subject . . . . significant for future development."

The first chapter is mainly devoted to a qualitative description of the features of turbulent flow in general, and to analogies between turbulence and the random motion of molecules. Then follows a concise description of the observations of turbulent flow in the shallow layer of atmosphere accessible to earthbound instruments and of the vertical gradients and diurnal variations of mean quantities (wind, temperature, etc.) affected by and effecting the turbulence. In four remaining chapters the author develops a mathematical treatment of turbulent diffusion, based in turn on empirical K (eddy viscosity) theory, on Prandtl mixing-length theory, and on the author's own empirical formulation of the Taylor correlation function, the function which measures the rate at which an eddying mass of air loses its identity during motion through its environment.

The older K-theory is rather quickly dismissed as sterile, but mixing-length and correlation-function theory are used to erect a seemingly substantial structure of equations for the transfer and concentrations of momentum, heat and matter from various sources. The structure has, however, little cement between the members or in the foundations. The cement one looks for is satisfying physical argument or idea but whether it will ever be applied to this particular edifice is questionable. The formulae can undoubtedly be useful in practical application but their range of validity is strictly within the compass of their test in the field. This is roughly the present state of the subject and to say so is no criticism of Professor Sutton's matter, though one might have expected that he would have stressed more strongly the quite empirical nature of the work. It would be little exaggeration to say that practically the whole of the theory presented in this monograph, Taylor's theorem of diffusion apart, is derivable from dimensional analysis and while he would be a poor physicist who spurned such aid, dimensional analysis is not enough to provide a very lively body of knowledge. The author indicates that ideas for a more or less fresh approach are beginning to emerge but, as he says, the ideas are almost still confined to the fluid-motion laboratory and can hardly yet face the atmosphere with its complicating density gradients and larger scales of motion.

Four-fifths of the earth's surface being water, one would have liked to find some reference in a treatment of terrestrial boundary-layer turbulence to the lowest 100 metres or so over the sea, for this boundary undoubtedly has special properties. Perhaps, when a second edition is called for, the author will enlarge his treatment in this sense.

The book is well printed, referenced and indexed, and will undoubtedly find a welcome as a most competent survey of a useful, if not very meaningful, phase of development in this entrancing but tantalizing subject. P. A. S.

Practical Microscopy, by L. C. Martin and B. K. Johnson. Pp. 124. Blackie's "Technique" Series. 2nd Edition. (London and Glasgow: Blackie and Son, Ltd., 1949.) 6*s*. 6*d*.

Any book which sets out to describe the technique of modern microscopy in 124 small pages must of necessity be incomplete in some respects. Let it be said at once that "Martin and Johnson" is one of the best books of its kind. It explains the basic principles in a clear and simple manner, yet manages to deal with many small points which are not covered by much larger volumes. The practical microscopist will find of particular value the many numerical examples to illustrate common problems, and the chapter on optical artefacts is one which cannot be read too often. It is easy to criticize omissions in a book of this size. One cannot help expressing some disappointment that a rather more thorough revision has not been undertaken. True, a chapter on the electron microscope and a section on phase contrast illumination have been added, but apart from this the book has scarcely been altered since 1931. Thus it is still stated that "the use of the polarizing microscope in biological work does not appear to have received the attention it deserves", whereas the literature on the subject now runs into thousands of papers, and the instrument itself has been vastly improved. The basic principles of ultra-violet microscopy are well described, but modern developments, in which one of the authors has played an important part, are not referred to.

Despite these criticisms, which the reviewer hopes will be remedied in the many future editions which the book deserves, "Martin and Johnson" remains a book which should be available to every microscopist. R. BARER.

Ionization Chambers and Counters: Experimental Techniques, by B. B. Rossi and H. H. Staub. Pp. xviii + 243. 1st Edition. (New York, Toronto, London: McGraw-Hill Book Co. Inc., 1949). 19*s*. 6*d*.

This volume is the second of the National Nuclear Energy Series (Los Alamos Project, Division V) and it has an obviously close association with the first—*Electronics* by W. C. Elmore and M. L. Sands. The first four chapters deal with the fundamental features of ionization and the general properties of detectors based upon the ionization process. Many readers will find Chapter 1 of special interest since it collects much useful data on the drift velocity and agitation energy of electrons; a good deal of material due to the authors themselves appears here for the first time and it refers closely to counting techniques. Chapters 2 and 3 are concerned with the general design considerations affecting the performance of chambers and the latter deals in very thorough fashion with the relatively new problem of the measurement of variable ionization. In Chapter 4 the basic ideas of the gas multiplication process are rather briefly presented. A series of curves gives the values of multiplication observed in a variety of gases in typical operating conditions. The studies of pulse shape, end-effect of tubes and multiple-wire arrangements follow. While the allocation of space to the different sections seems somewhat strange in a textbook all of it is of interest to the specialist in this field.

The last five chapters consider in some detail the design of detectors for beta-, gamma- and x-radiations (Chapter 5), alpha-particles, neutrons and fission products. Chapter 5 is comparatively modest in its scope and rather unbalanced. On the other hand, the compact discussion of α-ray spectroscopy is very meritorious. The two long chapters devoted to the neutron form the core of this section of the book; the whole subject is treated with originality and the authors are to be congratulated on their thorough treatment of the many difficulties encountered in this field. Again, a considerable fraction of the material is original. Finally, Chapter 10 discusses fission detectors.

The book suffers in several directions due to its origin and it is particularly disappointing that there is no bibliography. It is a notable contribution to the very limited number of texts on a subject of major importance in atomic and nuclear physics. The number of engineering drawings is unusually generous. All closely interested in the subject will be grateful to the authors for this excellent book. S. C. CURRAN.

Differential Equations, by HARRY W. REDDICK. Pp. x+288. 2nd Edition. (New York: John Wiley and Sons; London: Chapman and Hall, Ltd., 1949.) 24*s*.

This is a relatively elementary textbook, not, as its title implies, on differential equations generally, but on ordinary (as opposed to partial) equations. It could certainly be understood and used immediately after a first course in calculus, and might be useful to those who intend to be engineers or applied scientists, but not to those who intend to take their mathematical physics to any high standard. For the latter, the approach is much too lacking in rigour, a feature which is illustrated in the first chapter, where the difficulties associated with deciding how many arbitrary constants must appear in the solution of an equation of given order are glossed over. It is not a criticism of the book, but an indication of its scope, to say that existence theorems are lacking and that Sturm–Liouville systems are not mentioned, nor are their typical properties illustrated even in a simple particular case.

Within its restricted aims, however, the book is clear and accurate, and has the excellent feature of containing many good examples, with answers provided. It should serve well to teach the tricks of integrating those equations which have solutions expressible in closed form or in power series.

J. H. A.

Formulas and Theorems for the Special Functions of Mathematical Physics, by WILHELM MAGNUS and FRITZ OBERHETTINGER; translated by JOHN WERMER. Pp. 172. (New York: Chelsea Publishing Co., 1949.) No price.

The special functions of mathematical physics are those satisfying a certain number of differential equations familiar in different parts of theoretical physics, and include Bessel functions, spherical harmonics, hypergeometric functions and elliptic functions. As specialized cases, there are Laguerre functions, Hermite functions and many others.

In this book, which is intended as a reference work and not as a textbook, the main (and a good many rather recondite) properties of a great number of these functions are set down in convenient form. There cannot be many of the properties given for the cylindrical functions which are not to be found in Watson's treatise, but it is probable that they could be found rather more easily in this compendium than in the larger work.

The standard of mathematical rigour is high, and the gamma function, for example, is defined in a way which applies to all (real or imaginary) values of its argument.

The book gains a good deal in value by containing a chapter (not strictly justified by the title) on integral transformations and one on coordinate transformations which, rather unexpectedly, does not refer to group theory.

J. H. A.

Détermination d'un etat plan des contraintes à l'aide d'un extensiomètre à résistance electrique à trois directions (*rosette*): *Abaques Pratiques d'Emploi*, par F. T. SALLES. Pp. 41. (Paris: Publications Scientifiques et Techniques du Ministère de l'Air, No. B.S.T. 112, 1949.) 350 fr.

The use of resistor strain gauges for the measurement of stresses and strains in structures, etc., has now become common practice. If the stress-system is two-dimensional, it is necessary to measure the strains in three directions at a given point in order to determine the strains (or stresses) completely; one method of doing this is to use a rosette gauge, consisting of three independent, intersecting resistor gauges, cemented to the structure at the point in question. The reduction of the experimental results is tedious and, in this report, tables and nomograms are provided to facilitate this work when the three gauges of the rosettes are inclined at 45° or at 120° to each other (rectangular or equilateral rosettes, respectively). The publication should be very useful to workers who use the rosette type of gauge.

R. M. D.

CONTENTS FOR SECTION A

PAGE

PAGE
Prof. W. **Franz**. On the Theory of Diffraction 925
Miss M. E. **Pillow**. Intensity Distribution in Band-Systems of O_2 and O_2^+ . 940
Dr. S. P. **Sinha**. Ultra-Violet Bands of K_2 952
Mr. E. B. **Andrews** and Dr. R. F. **Barrow**. A New Ultra-Violet Band-system of GeBr 957
Mr. A. **Hedgran**, Dr. K. **Siegbahn** and Dr. N. **Svartholm**. A Large β-Spectrometer with Two-Directional Focusing for Precise Measurements of Nuclear Radiation 960
Dr. D. K. **Butt**. Experiments with a Magnetic Lens Spectrometer, employing Post-focusing Acceleration : the L-Auger Lines of Thorium-active Deposit . 986
Mr. C. H. **Collie**, Dr. H. **Halban** and Dr. R. **Wilson**. Photoelectric Dissociation of the Deuteron 994
Prof. L. **Jánossy**. On the Lateral Spread of Extensive Air Showers . . . 1009
Mr. **Vachaspati**. On the Use of β-Formalism of the Meson Field for Nuclear Interactions 1015
Dr. **Kai-Chia Cheng**. A New Method for Determining the Radial Distribution Function 1028
Mr. J. **Hamilton**. Statistical Equilibrium and Radiation Damping . . . 1036

Letters to the Editor :

Mr. E. **Rabinowicz**. A Photographic Method for determining Half-Lives . 1040
Dr. C. G. B. **Garrett**. An Order–Disorder Transition Curve in Cobalt Ammonium Sulphate below 0·1° K. 1042
Mr. E. W. **Fuller**. The Half-Life of the Isomeric State in ^{169}Tm . . 1044
Dr. J. B. **Birks**. Scintillations from Naphthalene–Anthracene Crystals . 1044
Reviews of Books 1046
Corrigendum 1050
Contents for Section B 1050
Abstracts for Section B 1051

ABSTRACTS FOR SECTION A

On the Theory of Diffraction, by W. **Franz**.

ABSTRACT. A new method is described for calculating the diffraction of an acoustical or electromagnetic wave by successive approximations. Kirchhoff's theory is a special case of the first step in the new method, which, unlike Kirchhoff's, is not restricted to black screens, and applies to long waves as well as short ones.

Intensity Distribution in Band-systems of O_2 and O_2^+, by M. E. **Pillow**.

ABSTRACT. The intensity distributions for three oxygen band-systems, the Schumann–Runge and Herzberg systems of O_2 and the second negative system of O_2^+, are calculated, and compared where possible with experimental results. Some suggested identifications of bands observed in emission from the night sky are checked by comparison with the calculations.

Ultra-Violet Bands of K_2, by S. P. SINHA.

ABSTRACT. Bands of K_2 have been photographed in absorption in the region λ 3924–λ 3420 A. on a Hilger E1 quartz spectrograph. The bands lie in two groups, one between λ 3924 and λ 3686 A. and the other between λ 3640 and λ 3420 A., and appear to belong to two different systems.

The bands in the first group have been analysed. They can be represented by the equation :

$$\nu=26493\cdot0+60\cdot6u'-0\cdot15u'^2-92\cdot64u''+0\cdot354u''^2,$$

where $u=v+\frac{1}{2}$.

The upper state of this system is considered to dissociate into 4^2S+5^2P atoms of potassium.

No analysis for the bands lying between λ 3640 and λ 3420 A. has been possible.

A New Ultra-Violet Band-System of GeBr, by E. B. ANDREWS and R. F. BARROW.

ABSTRACT. The vibrational analysis of a new system of GeBr has been made. The constants derived in cm^{-1} are :

	ν_e	ω_e	$x_e\omega_e$
$^2\Delta$ or $^2\Pi$	27252	$197\cdot_3$	$7\cdot_{18}$
	27156	$190\cdot_1$	$7\cdot_{25}$
$^2\Pi_{3/2}$	1150		
$^2\Pi_{1/2}$	0	$295\cdot_4$	$0\cdot7_2$

The upper state, probably $^2\Delta$ or $^2\Pi$, appears to converge rather rapidly to a limit at about 3·5 ev. above $v''=0$ of $^2\Pi_{1/2}$.

A Large β-Spectrometer with Two-Directional Focusing for Precise Measurements of Nuclear Radiation, by A. HEDGRAN, K. SIEGBAHN and N. SVARTHOLM.

ABSTRACT. In the present paper a large double focusing spectrometer is described, which is convenient for precise measurements of β- and γ-radiation. The magnetic field form is discussed and also the arrangement to measure the value of the field continuously with an accuracy appropriate to investigations of this kind. The difficulties due to electron straggling in the β-source or converter are emphasized, and this effect is for the moment the factor which prevents the full utilization of the high resolving power of the apparatus. A number of typical spectra are given to illustrate the applicability of the instrument to different cases.

Experiments with a Magnetic Lens Spectrometer, employing Post-focusing Acceleration : the L-Auger lines of Thorium-active Deposit, by D. K. BUTT.

ABSTRACT. A post-focusing electron accelerator has been developed to reduce the absorption of low-energy electrons by the Geiger counter window of a β-ray lens spectrometer. Its useful range is for β-rays having energies between 0 and 30 kev., and although, in general, for a portion of this range only partial transmission will be obtained, correction curves for the absorption can be accurately determined.

In order to test the spectrometer, an investigation of the L-Auger lines of thorium-active deposit has been made. An identification of the lines has been attempted.

Photoelectric Dissociation of the Deuteron, by C. H. COLLIE, H. HALBAN and R. WILSON.

ABSTRACT. The cross section for the photo-disintegration of the deuteron has been measured by detecting the photo-protons formed in a high pressure ionization chamber. The values obtained are $13\cdot91\pm1\times10^{-28}$ cm² for 2·62 MEV. γ-rays from radio-thorium and $15\cdot6\pm1\times10^{-28}$ cm² for the 2·76 MEV. γ-rays from radio-sodium.

On the Lateral Spread of Extensive Air Showers, by L. JÁNOSSY.

ABSTRACT. The lateral and angular spread of extensive showers is calculated in an approximation where ionization loss and incomplete screening are neglected (approximation A). The atmosphere is taken to be homogeneous. It is shown that the spread at the cascade maximum is equal to the average spread, but the spread increases considerably throughout the shower; the above results are in qualitative agreement with results given by Borsellino. The quantitative discrepancy between the results of Borsellino and other authors is commented on.

On the Use of β-Formalism of the Meson Field for Nuclear Interactions, by VACHASPATI.

ABSTRACT. Following a method due to Harish Chandra, a correspondence between the Proca and the β-formalisms of the meson field theory is developed. An important feature of the method is that one can at any stage of calculations pass over from one formalism to the other. No explicit representation of the β-matrices is required nor need one know the exact relationship between the elements of the wave function ψ and the Proca field quantities U_μ and $G_{\mu\nu}$. This introduces considerable simplification in the introduction of nuclear interaction as compared with earlier treatments where one has to choose a special explicit representation of the β_μ's and identify the various elements of the matrix ψ with the components of U_μ and $G_{\mu\nu}$. The nuclear interaction is, in fact, here introduced very simply in an elegant way. The quantization is performed for the vector, scalar and pseudoscalar cases assuming in all cases both charge and dipole types of couplings. As an illustration of the method we calculate the scattering of positively charged mesons by protons. The spur calculations, so tedious in the earlier works, are now replaced by simple relations which are easily handled, and the results are obtained with about as much labour as is used in Proca's case with the additional advantage that both the longitudinal and the transverse parts of the meson field can here be treated together.

A New Method for Determining the Radial Distribution Function, by KAI-CHIA CHENG.

ABSTRACT. It is shown in this paper that Green's integral equation resulting from the Kirkwood approximation for the radial distribution function can be transformed into a differential equation of infinite order, to which the method of successive approximation can be applied without ambiguity. The solution at moderately high temperatures is discussed. The result is a correction to the usual Boltzmann distribution function; the radial distribution function oscillates with decreasing amplitudes at small distances and tends to unity at large distances. The validity of the approximations is discussed in two appendices.

Statistical Equilibrium and Radiation Damping, by J. HAMILTON.

ABSTRACT. The reason why the usual form of radiation damping theory does not satisfy the principle of detailed balancing is discussed. A form of damping theory satisfying statistical equilibrium is developed.